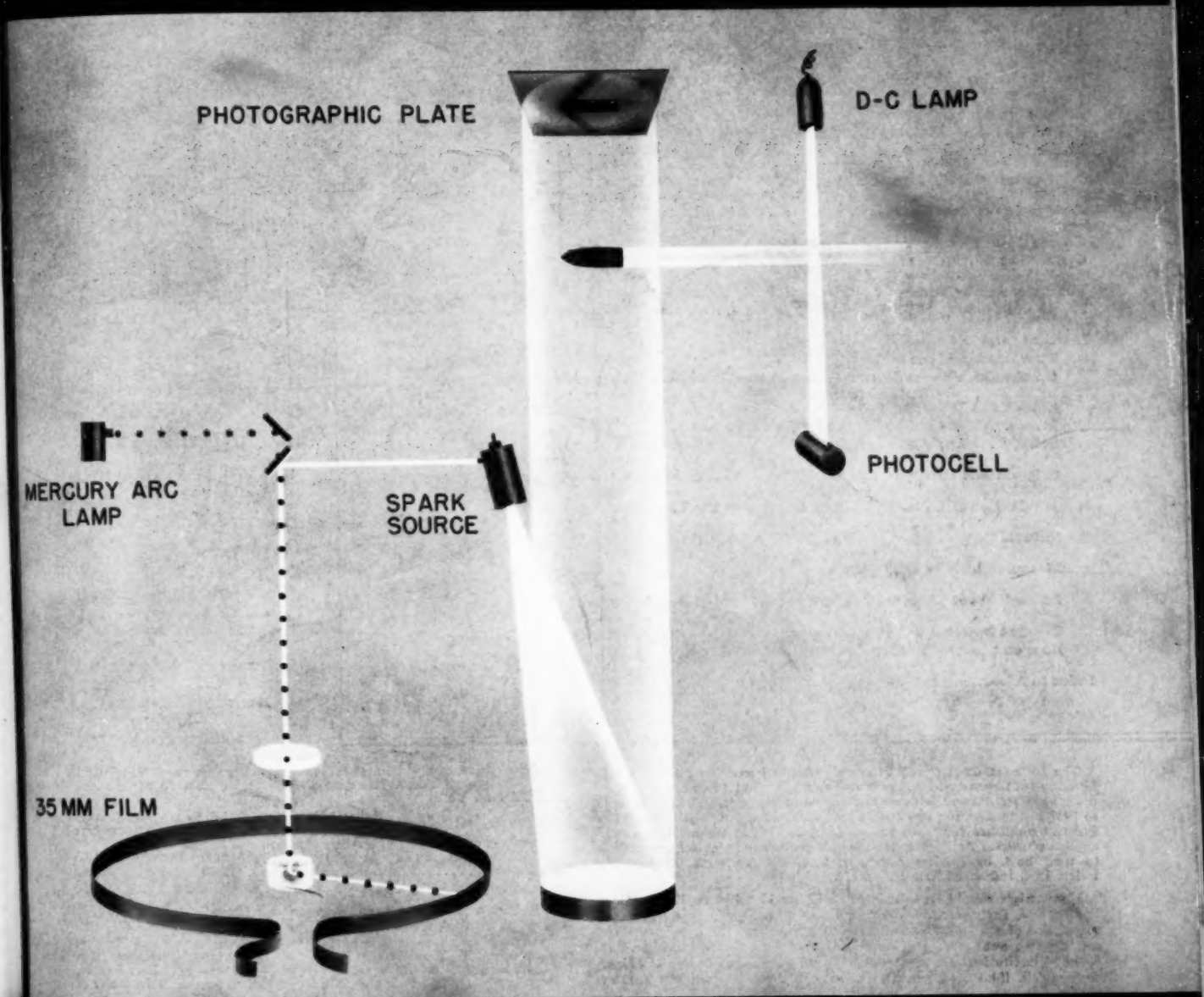
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 21, No. 11 NOVEMBER, 1950

In This Issue

		Page
On Crazing of Linear High Polymers	C. C. Hsiao and J. A. Sauer	1071
End-Cooling of Power Tube Filaments	John W. Clark and Ralph E. Neuber	1084
Space-Charge Effects in Electron Beams and Their Reduction by Positive Ion Trapping	E. G. Linder and K. G. Hernqvist	1088
Stress-Strain Properties of Natural Rubber under Biaxial Strain	B. B. S. T. Boonstra	1098
On the Endurance Limit of a Round Bar with Longitudinal Grooves	H. Ōkubo	1105
Magnetic Field of a Plane Circular Loop	C. L. Bartberger	1108
Effect of Coating Composition of Oxide-Coated Cathodes on Electron Emission	E. G. Widell and R. A. Hellar	1115
Surface Waves and Their Application to Transmission Lines	Georg Goubau	1119
Note on Control Area	Thomas M. Stout	1129
Small Electron Microscope	John H. Reisner and Edmund G. Dornfeld	1131
Artificial Field Equations for a Region where μ and ε Vary with Position	P. D. P. Smith	1140
Hypersonic Research Facilities at the Ames Aeronautical Laboratory	Victor I. Stevens	1150
Application of Rate-Process Theory to Glass. I. Breaking Strength	F. B. Hodgdon, D. A. Stuart, and F. E. Bjorklund	1156
Application of Rate-Process Theory to Glass. II. Viscosity	F. B. Hodgdon and D. A. Stuart	1160
Circuit Analysis of Linear Varying-Parameter Networks	Lotfi A. Zadeh	1171
Convection Currents in Porous Media. III. Extended Theory of the Critical Gradient	F. T. Rogers, Jr. and H. L. Morrison	1177
Theory of Axially Slitted Circular and Elliptic Cylinder Antennas	D. R. Rhodes	1181
Frequency Factor and Activation Energy for the Volume Diffusion of Metals	G. J. Dienes	1189
Letters to the Editor:		
Thorium Sulfide as a Thermionic Emitter	T. E. Hanley	1193
On the Fredrickson-Eyring Theory of the Mechanical Behavior of Metals	J. Fleeman and G. J. Dienes	1193
Base-Metal Effects in Thoria-Coated Filaments	Herbert Nelson	1194
Erratum: The Mechanical Properties of Glass	F. W. Preston	1195
Scheme for Analyzing Capillary Measurements on Non-Newtonian Liquids	Orrin H. Clark and M. L. Deutsch	1195
Original Authorship of Formulas Given in Clark and Deutsch's Paper	M. Reiner and G. W. Scott Blair	1195
Application of Liouville's Approximation to the Blind Navigation Problem	J. J. Gilvarry and S. H. Browne	1195
Natural Crystal Greenockite (Cadmium Sulfide) as an X-Ray Detector	E. L. Criscuolo and D. T. O'Connor	1196
Measurements on the Diffusion of Interstitial Atoms in B.C.C. Lattices	C. A. Wert	1196
Remark about Standardization and "A Dimensional Analysis of Metal Cutting"	André Martinot-Lagarde	1197
Tearline Patterns in Ferrochromium	C. A. Zapffe and F. K. Landgraf	1197
Note on "Interference of Growing Spherical Precipitate Particles"	Hershel Markovitz	1198
Index to Advertisers		xxxii

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Earle C. Gregg, *Assistant Editor*

Associate Editors

Howard H. Aiken
James Hillier
W. James Lyons
Philip Morrison
Eric Reissner
Raymond J. Seeger
Frederick Seitz
Charles S. Smith

Publication of the

American Institute of Physics

OFFICERS

George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*

GOVERNING BOARD

J. W. Beams
C. Paul Boner
Wallace R. Brode
W. F. Fair, Jr.
J. B. Fisk
G. P. Harnwell
George R. Harrison
W. V. Houston
Rudolph Kingslake
Paul Kirkpatrick
Paul E. Klopsteg
F. W. Loomis
Philip M. Morse
J. R. Oppenheimer
George B. Pegram
Harold K. Schilling
Francis G. Slack
John C. Steinberg

Consultant on Printing

Melvin Loos

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Assistant Treasurer and Comptroller*
Ruth F. Bryans, *Publication Manager*
Betty H. Goodfriend, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members and associate members of the American Institute of Physics	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 21, Number 11 November, 1950

On Crazing of Linear High Polymers*

C. C. Hsiao and J. A. Sauer
The Pennsylvania State College, State College, Pennsylvania
(Received April 21, 1950)

The effect of various variables, such as type of stress, stress magnitude, duration of stressing, and environment on the initiation and development of crazing in linear polymers is discussed. The basic nature of crazing is investigated in some detail for polystyrene specimens by means of the light microscope, the electron microscope and the x-ray spectrometer. The results of these observations and their bearing on the fundamental group structure of polystyrene molecules is presented.

The relationship between crazing and orientation is discussed, as well as the effect of both of these factors on the mechanical properties. The experimental results are compared, wherever possible, with the previously reported data of other investigators. A short discussion is then given of a theory of crazing from the point of view of the molecular structure of the material.

INTRODUCTION

THE phenomenon of crazing of thermoplastic materials, especially polystyrene, has been under intensive investigation recently by many scientists.[1–5] One of the aims of this research is to discover the cause of crazing and to find methods of eliminating it, if possible. The inception and subsequent growth of crazing are undesirable features that tend to destroy optical transparency and influence the mechanical behavior. Although considerable information is now available on the macroscopic effects of crazing and their relation with other variables such as birefringence, tensile stress, temperature, etc., the basic nature of crazing from the viewpoint of the molecular structure of the material is still not very well understood. It seems advisable, therefore, not only to investigate and record the various external manifestations of this general phenomenon but to examine critically the fundamental nature of crazing from the microscopic and submicroscopic viewpoint. This is the intent of the present investigation which reports the results of observations made with the light microscope, the electron microscope and the x-ray spectrometer. It is hoped that the results of this investigation will aid in providing a better understanding of the crazing of linear high polymers in general.

DEFINITION OF CRAZING AND GENERAL OBSERVATIONS

By "crazing" of thermoplastic materials is meant the development of a state in which the normal optical transparency of the material is reduced. For such thermoplastics as polystyrene or polymethyl methacrylate, crazing may be described in terms of the development of mechanical cracks visible to the naked eye[1] or as a blushing of the material produced by the scattering and reflection of light.[5] Several different patterns of crazing have been observed. Russell[1] has classified them as stress crazing, solvent crazing and combinations of the two. Maxwell and Rahm[4] have likewise reported observations of crazing produced both by application of stress and by solvent action. Sauer, Marin and Hsiao[5] have shown pictures of stress crazing occurring on the fracture surfaces of tension and bending

* This invited paper was presented at the meeting of the Division of High Polymer Physics of the American Physical Society in New York City, February 4, 1950.

[1] E. W. Russell, *Studies on Polymethyl Methacrylate, Part III—Crazing Effects*, Report No. Chem. 447, Royal Aircraft Establishment, Farnborough Hants, England (August, 1948). Nature, London, No. 4186, **165**, 91 (1950).

[2] B. Maxwell and L. F. Rahm, Ind. Eng. Chem. **41**, 1988 (1949).

[3] J. R. McLoughlin, *Crazing of Polystyrene Films*, Plastics Laboratory Technical Report 12B, Princeton University, Princeton, New Jersey (December 15, 1948).

[4] B. Maxwell and L. F. Rahm, *Factors Affecting the Crazing of Polystyrene*, Plastics Laboratory Technical Report 14B, Princeton University, Princeton, New Jersey (May 5, 1949).

[5] J. A. Sauer, J. Marin, and C. C. Hsiao, J. App. Phys. **20**, 507 (1949).

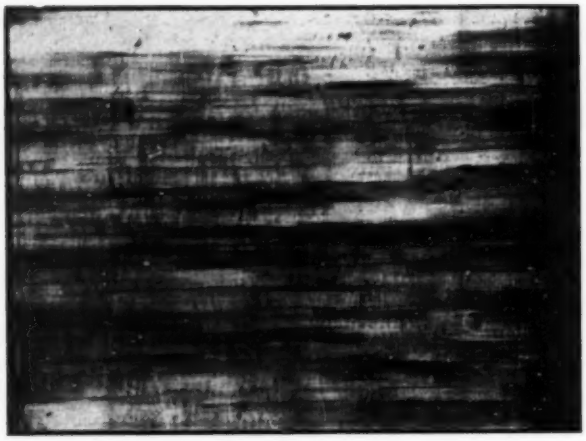


FIG. 1. Crazing cracks in polystyrene.

creep specimens of polystyrene subject to different stresses and different durations of load application. The latter authors have also summarized[5] their observations of the phenomena of stress crazing and interpreted the phenomena observed in terms of changes of molecular structure. It was their feeling that crazing is not merely a mechanical crack phenomenon, although it may be and frequently is closely associated with visible mechanical cracks.

In order to differentiate between these somewhat different descriptions of essentially the same phenomenon, we will consider "crazing cracks" to be the visible cracks produced only after "crazing," in the form of real but submicroscopic discontinuities, has previously come into existence. Crazing may then refer to or be observed in the form of a blushing or haze of a previously transparent material. In other words, crazing is not necessarily actual cracks, but may be some sort of infinitesimal openings which are either still within the field of molecular attraction or are physically prevented by other neighboring molecules from developing into cracks.

The most basic fact about the development of crazing in specimens of linear high polymers subject to load is that the observed crazing cracks are always found to be at right angles to the direction of maximum tensile stress. A microphotograph indicating this behavior for polystyrene is shown in Fig. 1. If, however, the specimen is subject to pure compression rather than tension, no crazing will occur.

Both of the above facts are simultaneously observable when a rectangular specimen is subject to pure bending, as under this type of stress condition one-half of the specimen is subject to tension and the other half to compression. Figure 2 shows two views of a polystyrene specimen which had been continuously subject to a constant bending moment until fracture occurred. The left view clearly shows the absence of crazing on the compression side of the specimen and its strong presence on the tension side. The right view showing the fracture surface reveals on close examination two distinct types of cracks; one type is the fracture cracks which spread radially outward from a more or less point source usually coincident with a flaw, and the second type is the crazing cracks which are seen as normal openings or planes starting from the surfaces of the specimen.

Increase of crazing of polystyrene specimens with increase of time of load application is shown by the six specimens of Fig. 3, each of which was subject to a constant tensile stress varying from 3250 p.s.i. for the specimen on the left to 5500 p.s.i. for the specimen on the right as indicated. The specimens were allowed to remain under these stresses until fracture occurred; for the 5000-p.s.i. specimen the duration of test was very short (minutes), while for the 3250-p.s.i. specimen hundreds of hours elapsed before the specimen failed.

The size, density and depth of penetration of crazing openings depends not only on stress and time but also on the particular material under investigation. For example, it is found that for comparable stress and time conditions, the crazing cracks on the surface of polymethyl methacrylate specimens are usually larger and less dense than those in polystyrene specimens, but at the same time, do not extend inward from the surface as deeply. This latter fact may be due to the

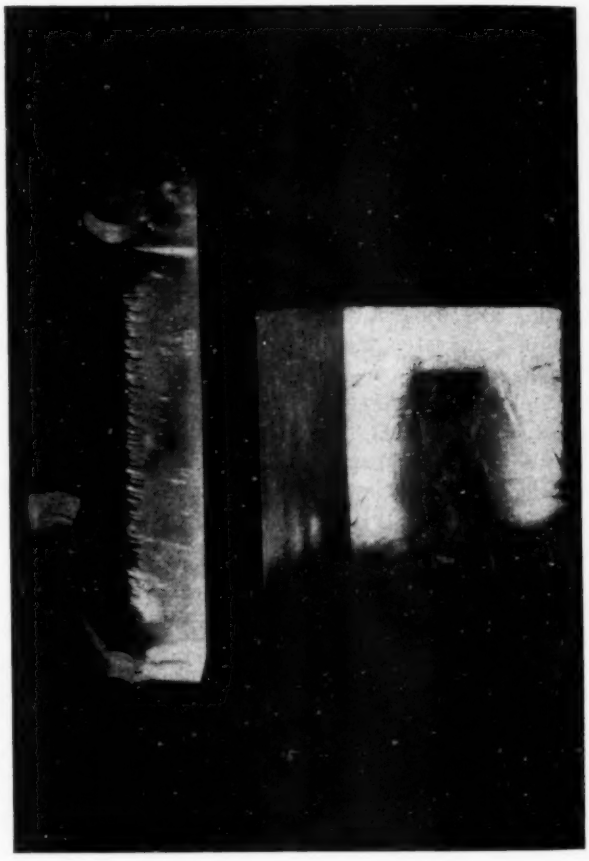


FIG. 2. Bending creep specimen of polystyrene showing crazing along length and on cross section of tension side of specimen.

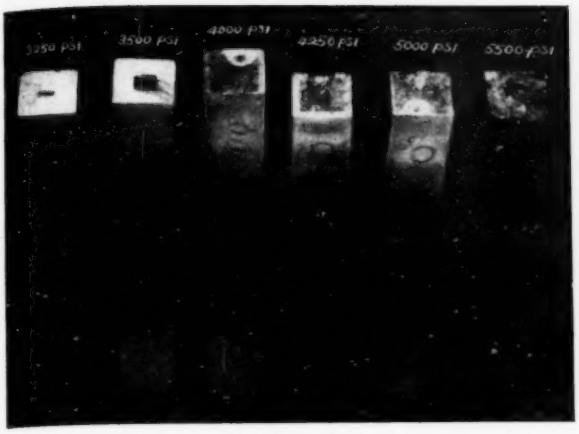


FIG. 3. Fracture surfaces of crazed tension creep specimens of polystyrene subject to different stresses and different durations of load application.

inability of the relatively "brittle" polystyrene specimens to relieve themselves of high local stress concentrations as readily as the more ductile methacrylate specimens.

Another rather interesting observation is that crazing may extend throughout the entire cross-sectional area and yet the specimen can continue to carry load even though reduced in value. This is well shown by the stress-strain curve in Fig. 4 along with that of a standard amorphous polystyrene specimen. While the stress-strain curve of the crazed specimen indicates that it is considerably weaker than the original specimen, it is clear that even the entirely crazed regions of a specimen can carry considerable stress despite the existence of the numerous crazing openings or cracks. The stress-strain curve suggests that the material has been locally oriented due to stretching as the modulus of elasticity of the earlier part of the curve appears to be slightly higher than that of the amorphous material. Above 1500 p.s.i., the crazing openings grow rather rapidly into visible cracks and fracture of the specimen occurs at usual strain values but at lower stress values.

In order to determine whether any similar damage had been done to the apparently non-crazed central portions of a tension creep specimen, the crazed outer portion of the polystyrene specimen was machined off and the remaining central core then subject to tension. The stress-strain curve is shown in Fig. 4 and, on comparison with that of the original amorphous specimen, the change, if any, is seen to be very slight. The modulus of elasticity is slightly higher suggesting that some short-range orientation of the polymer chains might have occurred. Also, a slightly lower ultimate strength might be partly due to the existence of submicroscopic crazing cracks too small to be seen. Both of these effects are, however, small enough to be caused by unavoidable differences between specimens.

In order to find whether compression would be beneficial in eliminating crazing previously produced by tension, a piece of a crazed polystyrene specimen was compressed and released three different times and then tested in tension. Although all the crazing cracks had seemingly disappeared after the third compression, they reopened after tension was applied. The four stress-strain curves are shown in Fig. 5 and the tension stress-strain curve is shown to a larger scale in Fig. 4. While it is clear that the material does not regain completely its original strength, it may be noted that its tensile strength is slightly higher than that of the fully crazed specimen, perhaps indicating that some of the submicroscopic crazing openings have been partially restored thereby providing some additional resistance to loading.

The most noticeable effects of the compressions were the gradual reduction and apparent elimination of the crazing planes and the large increase in ductility.

The effect of the environment on the initiation and development of crazing has been studied by running tensile creep tests on specimens of polystyrene subject

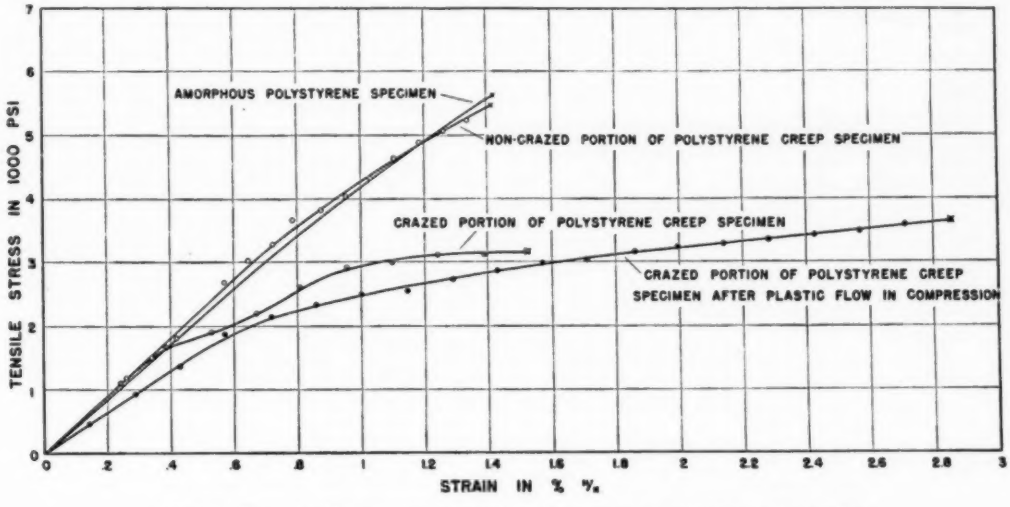


FIG. 4. Stress-strain curves of crazed and non-crazed polystyrene in tension.

to various atmospheric conditions. It was observed that in dry nitrogen, a specimen crazed rather easily and crazing seemed to penetrate deeper than in specimens subjected to the same stress in the ordinary atmosphere (50 percent R.H.). Less crazing was observed for a specimen tested in wet nitrogen. These results seem to indicate that water vapor has a healing effect on cracks that would otherwise open up under the prevalent stress system. This conclusion was subject to additional verification by running one creep test on a polystyrene specimen completely immersed in water. As expected, there was much less penetration of the crazing cracks into the specimen. It was also observed that crazing could be reduced considerably by coating the exterior surfaces of the polystyrene specimen before test. The results were somewhat spotty due probably to the fact that it was rather difficult to get a continuous uniform adherent coating on the square specimen. Nevertheless, the particular coating used, a compound of styrene monomer recommended for attachment of strain gauges to polystyrene specimens, was helpful in reducing crazing.

Crazing can also be produced by action of ultraviolet or x-ray radiation. The effect of the radiation is to cause depolymerization of the material. After the radiation is removed, the material will tend to repolymerize although a fairly long period of time may be required if the temperature is low. On repolymerization, the material will shrink, thereby inducing tensile stresses which in turn will cause crazing cracks to open up at right angles to the stress direction. Since repolymerization is at random, crazing patterns produced by ultraviolet radiation show randomly directed crazing openings or planes and hence are somewhat similar to crazing patterns produced by solvent action.

Although the external manifestations of crazing are now fairly well established for a wide variety of external conditions, little is known about their relation to internal structural changes of the material itself. If crazing could be directly related to elastic or plastic extension of the chain linkages or to tensile or shear deformation between individual chains or to other rearrangements of the polymer chain structure, a major step in general understanding of the phenomena would have been made and would unquestionably lead to fruitful suggestions concerning its elimination.

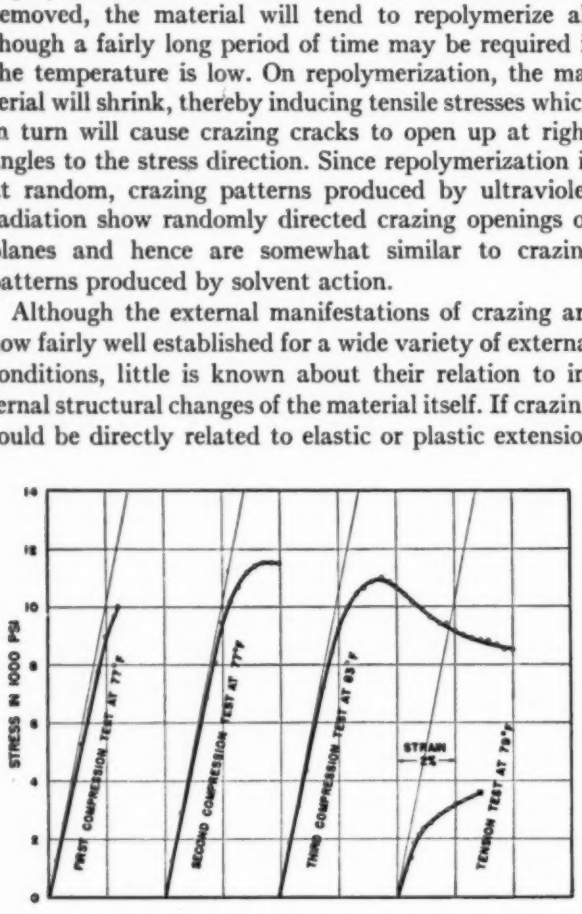

STRESS IN 1000 PSI
FIRST COMPRESSION TEST AT 77°F
SECOND COMPRESSION TEST AT 77°F
THIRD COMPRESSION TEST AT 83°F
STRAIN
TENSION TEST AT 75°F

FIG. 5. Effect of compression on the stress-strain curve of crazed polystyrene specimen in tension.

To accomplish the above, the first step would appear to be a careful detailed examination of the crazed and non-crazed surfaces by the best means at our disposal so that whatever changes were present could be detected. This examination is made first visually, then by the light microscope and finally by the electron microscope. In addition, a study has also been made of the crazed and non-crazed portions of the specimen using x-ray diffraction methods. The results of this general examination are now presented.

DETAILED DESCRIPTION OF A PARTIALLY CRAZED POLYSTYRENE SPECIMEN

Lustrex of an average molecular weight between 67,600 and 68,700 was used in this investigation. The stress-strain properties of this material have been described previously by the authors.[5]

The fracture surfaces of tension creep specimens of polystyrene (Fig. 3) subject to different stresses and different durations of load application reveal the dependence of crazing and its growth on the tensile stress magnitude and its duration. The fracture surface is seen to consist of a crazed and a non-crazed portion, the relative amounts of each depending on the magnitude and duration of stress to which each specimen was subjected. The amount of crazing is greater for the specimens subjected to the lower stress levels because the time of load application for these specimens was much greater. The relation between tensile stress and time to fracture appears to be a linear one on a semi-log plot (see Fig. 6).

The specimen at the extreme right of Fig. 3 was loaded at a stress approximately equal to the ultimate strength of polystyrene in tension (~5500 p.s.i.) and hence fractured shortly after application of the load. There is no apparent crazing and no crazing cracks are visible on the fractured surface. The second specimen from the right was stressed at 5000 p.s.i. and hence lasted for some time longer before it failed. The growth of crazing from the edges normally inward toward the center of the specimen is very evident. It may also be noted that the crazing regions are uniformly distributed throughout the total gauge length of the specimen. (This is especially evident when cross-sectional planes other than the fractured surface are examined.) In general, as we go from right to left in Fig. 3, penetration of crazing increases. Meanwhile, the duration of load application also increases in the manner of Fig. 6.

In order to clearly show the difference between crazed and non-crazed material, the fracture surface of a partially crazed specimen of polystyrene was photographed as shown in Fig. 7. This photograph was taken

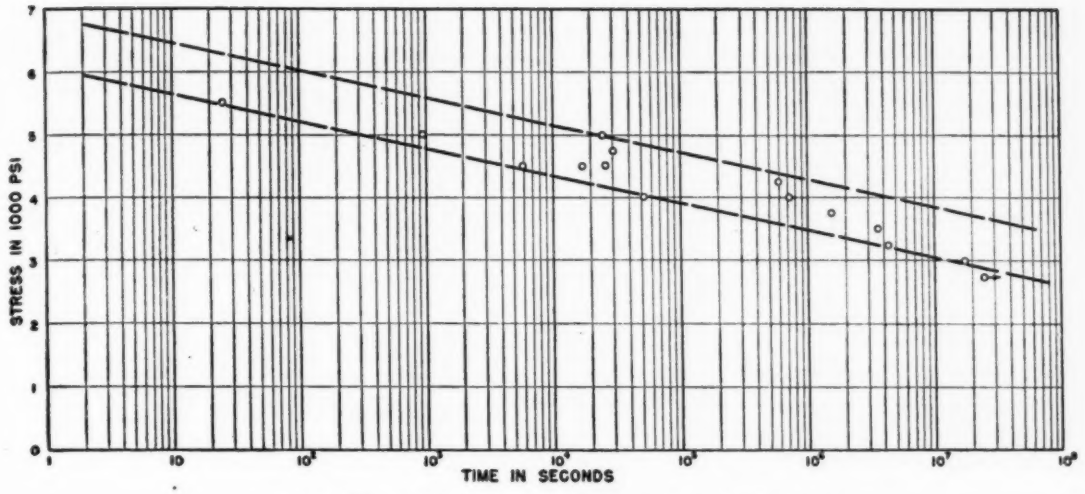


FIG. 6. Effect of tensile stress on fracture time of polystyrene.

by placing the polystyrene specimen against a black background so that the only reflected light would be that coming from the fracture surface. The crazed material along the edges of the specimen is easily differentiated from the centrally located non-crazed material by the smoother, mirror-like appearance.

It should be emphasized that when fracture does occur, the fracture cracks of the material are not the same as crazing cracks and that the source of fracture is usually some flaw in the material and not one of the crazing openings. In Fig. 7, this flaw is located at the center of the smooth circular region. After initiation, the fracture cracks proceed (probably initially at a constant rate) in directions at right angles to the principal stress direction. In Fig. 7, they are seen to extend more or less radially outward from the region in which the fracture first started whereas the crazing cracks are seen to extend normally inward from all edges of the $\frac{1}{2}$ in.$\times\frac{1}{2}$ in. specimen. During the propagation of the crazing and fracture cracks, the effective cross-sectional area of the specimen reduces and hence the tensile stress increases. Accordingly, the molecular rupturing stress is reached at an increasing rate near the neighborhood of the fracture surface. Possibly, the velocity of propagation of the rupture is also increased. Two clear concentric rings just outside the smooth circle indicate that this type of propagation at a varying rate has taken place. Similar phenomena on the fracture of glass have been studied and reported by Smekal.[6]

LIGHT MICROSCOPIC EXAMINATION OF CRAZED POLYSTYRENE

When the fracture surface of a polystyrene tension creep specimen is examined under higher magnifications, there are additional interesting features. Rochow and Rowe[7] have previously reported the characteristic structures of both fractured polystyrene and polymethyl methacrylate surfaces, but did not report on the nature of the crazing regions and surfaces. Under microscopic examination, the blushed or crazed region of Fig. 7 is seen to be entirely different from the amorphous, non-crazed central region. The crazing surfaces are found to be smooth parallel flat surfaces even when examined at a magnification as high as 500×. The flatness of these crazing surfaces as well as their parallel existence throughout the length of the specimens are largely responsible for the high reflectivity of these regions to white light.

Under a magnification of 50×, the fracture surface of a tension creep specimen of polystyrene shows many interesting patterns. Figure 8a is the upper inset of

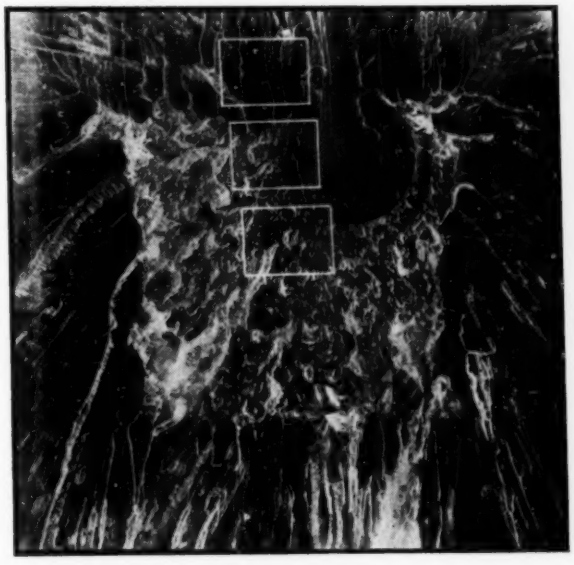


FIG. 7. Fracture surface of tension creep specimen of polystyrene showing crazed and non-crazed areas and a starting circle from which the cracks are radiated throughout the entire cross-sectional area.

[6] A. Smekal, Glastech. Ber. **13**, 141 (1935); Glastech. Ber. **13**, 222 (1935); J. Soc. Glass Tech. **20**, 432 (1936).
[7] T. G. Rochow and F. G. Rowe, Anal. Chem. **21**, 461 (1949).

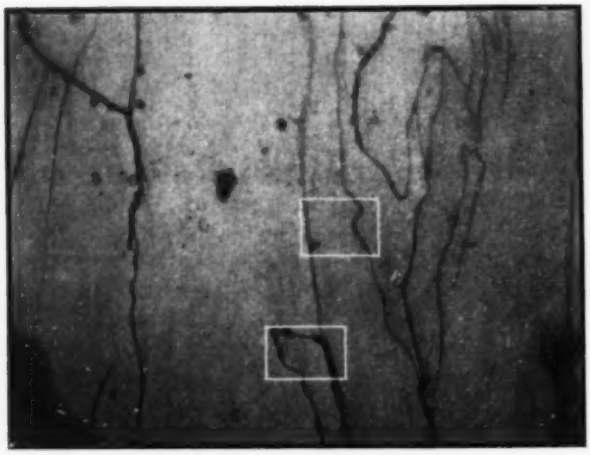


FIG. 8a. Fracture surface of tension creep specimen of polystyrene showing crazed areas and cracks (upper inset of Fig. 7).

FIG. 8b. Fracture surface of tension creep specimen of polystyrene showing the intersection of crazed and non-crazed sections and the starting circle (middle inset of Fig. 7).

FIG. 8c. Fracture surface of tension creep specimen of polystyrene showing the rings of propagation (lower inset of Fig. 7).

FIG. 9a. Fracture surface of tension creep specimen of polystyrene showing crazed areas and inner structure (upper inset of Fig. 8a).

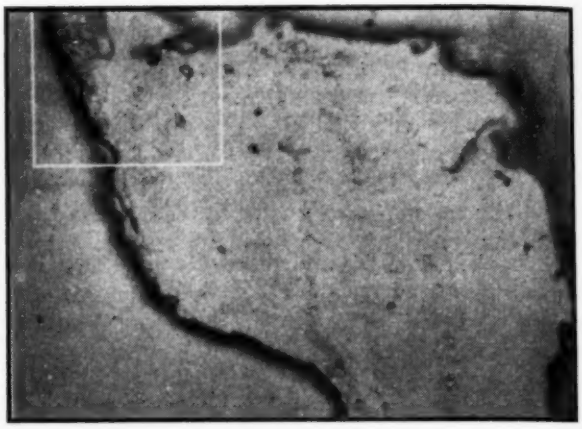


FIG. 9b. Fracture surface of tension creep specimen of polystyrene showing different elevations of crazed areas (lower inset of Fig. 8a).

Fig. 7 which illustrates principally many regions of crazing cracks. These surfaces are quite smooth and reflect white light, if not completely, at least, to a large degree. Both Russell[1] and Maxwell and Rahm[2,4] have made investigations of crazed high polymers with reflected light as the means of determination of the amount of crazing or crazing cracks. It is evident that practically the entire crazed section of a crazed polystyrene specimen can reflect light. The black lines in Fig. 8a are the boundaries of the crazing planes. These planes do not necessarily lie at one elevation as will be seen under higher magnifications later.

The termination of the crazed regions and their juncture with the non-crazed central portion of the specimen is shown in Fig. 8b. The upper white regions are the flat crazing areas separated from one another only by fracture or crazing cracks. The central rupture surface is no longer smooth. As a matter of fact, it is so much different in its elevations that a micrograph at 500× magnification will not show the details of even a very small portion of the texture. Figure 8b shows also,

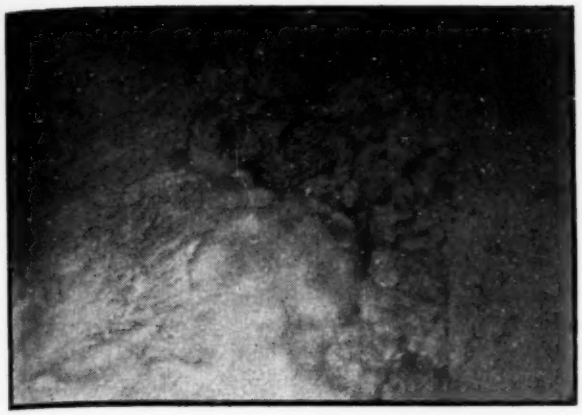


FIG. 10a. Fracture surface of tension creep specimen of polystyrene showing the connection of edge of the starting circle and the first ring of propagation (left inset of Fig. 8b).

FIG. 10b. Fracture surface of tension creep specimen of polystyrene showing the wavy colored bands occurring inside the starting circle (middle inset of Fig. 8b).

on the right half, part of the smooth starting circle which surrounds the flaw from which the fracture crack propagated. To the naked eye, this circular area is somewhat similar to the crazing area. However, it possesses an entirely different texture. It is a region having needle-like structure with no preferred orientation. Also, to be seen near this region are wavy and continuous lines or contours, indicating the presence of cavities. The clear arc in the middle of this figure is a part of the boundary of the starting circle. Successive rings of propagation are clearly shown in Fig. 8c. As was mentioned before, this is probably indicative of a changing rate of propagation of the crack that continues until failure of the total cross section of the specimen takes place.

The previous pictures have indicated that the surface of the crazing crack is different from that of the rupture crack originating from a flaw. This difference is even more strikingly illustrated by the higher magnification pictures of Figs. 9 and 10. Figures 9a and 9b two are pictures of the surfaces of the crazing areas. In these crazing planes, rather fine structure of the crazing surfaces shows that the smooth mirror-like surface is not a flat plane. Figure 10a shows the fracture surfaces in the region mostly bounded by the first ring of propagation. Figure 10b shows the enlarged cavity and the fine structure in the bottom of the cavity. In Fig. 10c the needle-like structure of the inside region of the starting circle is shown. This structure is somewhat similar to crazing patterns that are induced either by absorption of solvent vapor or by radiation depolymerization. Such patterns likewise show random orientation of the crazing openings. It should be noted that these cylindrical needles or openings are distributed at different elevations, whereas the fine structures on the surface of stress crazing cracks such as those shown in Figs. 9a and 9b are essentially located on one single plane. This latter fine structure is more clearly revealed by still higher magnifications such as Fig. 11 showing the inset of Fig. 9b.

ELECTRON MICROSCOPE EXAMINATION OF CRAZED POLYSTYRENE

In order to show the fine structure of crazed material still more clearly, electron micrographs were prepared. The results of both silica shadowed and beryllium shadowed crazed surfaces of polystyrene show similar structure. Here, only the electron micrograph of a beryllium shadowed crazed fracture surface is illustrated as shown in Fig. 12. This picture apparently correlates very well with Fig. 11. The double lines probably indicate the difference in elevations between two crazing planes. The order of the elevation in this particular picture is about two-tenths of a micron. The smoothness of the crazing surface is estimated to be less than one-tenth of a micron.

It is difficult to obtain a definite interpretation of this picture in terms of the polystyrene molecular structure. Are the small circles, for example, the end views of a group of parallel polymer chains? Under the same conditions of preparation, occasionally figures of the

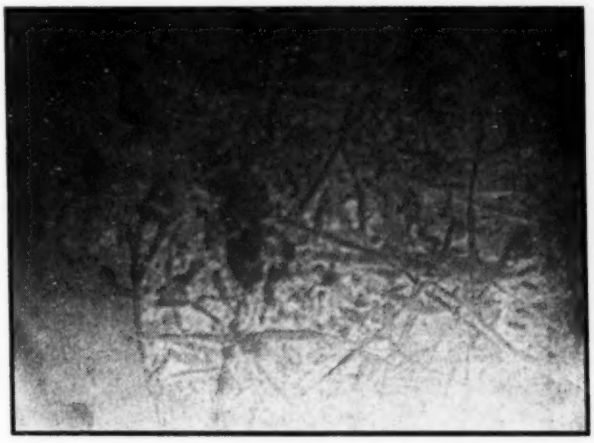


FIG. 10c. Fracture surface of tension creep specimen of polystyrene showing the needle-like structure occurring inside the starting circle (right inset of Fig. 8b).

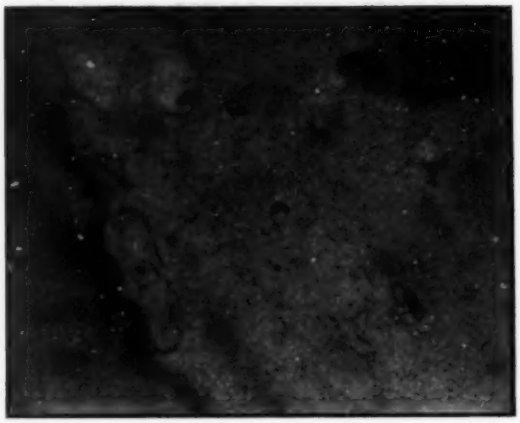


FIG. 11. Fracture surface of tension creep specimens of polystyrene showing the fine structure of crazed area at different elevations (inset of Fig. 9b).

type of Fig. 13 are obtained. Are these pictures of the inverse surface of Fig. 12 or do they indicate the existence of local regions or domains bounded by networks? In contrast to Fig. 12, showing a crazed surface of polystyrene, a molded surface of the same material was also recorded as shown in Fig. 14. This molded surface is perpendicular to the crazing surface. Many fine scratches are shown and the central breakage of the beryllium film may possibly indicate one of the actual large crazing cracks.

In addition to the electron-microscope pictures of the molded and fractured surfaces of solid polystyrene specimens, several additional electron-microscope pictures were taken of oriented polystyrene film. This was done both because the crazing properties of polystyrene are believed to be closely associated with orientation and also to provide some knowledge of the physical behavior of long-chain molecules under their natural configurations. The film was obtained by first dissolving the resin in a solvent and then allowing a drop of the solution to fall into water. As the drop comes back

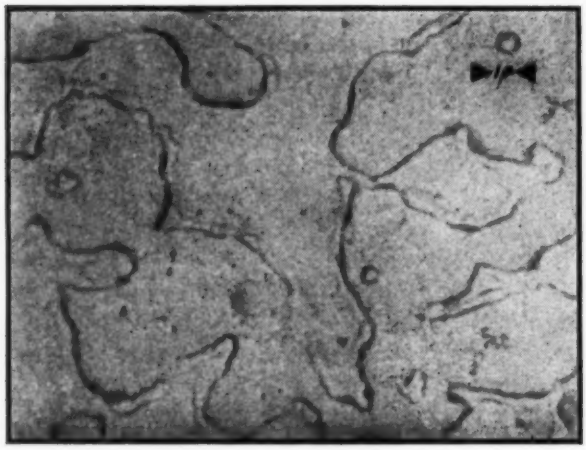


FIG. 12. Electron micrograph of beryllium shadowed crazed polystyrene showing crazing surface at various planes.

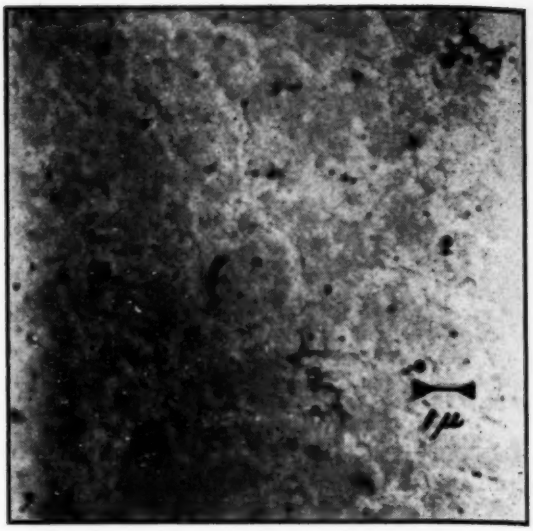


FIG. 13. Electron micrograph of beryllium shadowed crazed polystyrene showing crazing surface with various regions of domain structure.

to the water surface, it spreads out radially, the solvent evaporates and the plastic film is left behind on the surface of the water.

Figure 15 is an interesting picture of polystyrene film obtained by dropping polystyrene ethyl bromide solution on a cold water surface. This picture does seem to show the possible existence of domains bounded by networks and hence this type of structure may be the proper interpretation of Fig. 13. Since the formation of domains of polystyrene molecules is probably a very temperature dependent phenomenon, somewhat more striking results might be expected if the ethyl bromide polystyrene solution is dropped in hot water rather than cold. Such indeed is the case. Fig. 16 shows the electron micrograph of polystyrene film when prepared in this manner. This interesting picture reveals the

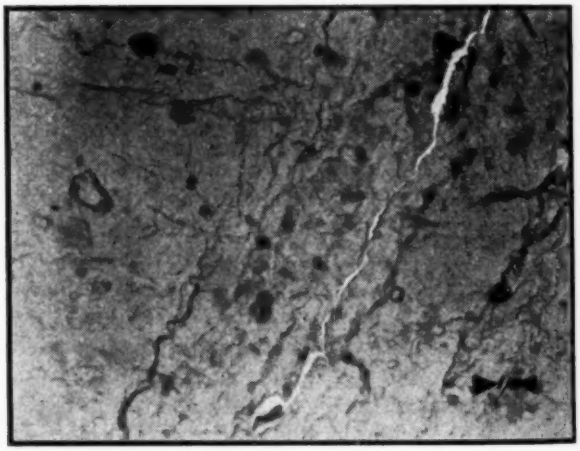


FIG. 14. Electron micrograph of beryllium shadowed polystyrene surface (perpendicular to crazed surface) showing various scratches on the molded surface.

tendency of the material to agglomerate in the form of rods or long slender domains. The picture shows two layers of rodlets both of which are oriented in one general direction. The proper interpretation of this picture is not yet well established but the cylindrical rods are quite possibly more or less parallel pencils of long chain molecules.

The technique employed in the preparation of an oriented film of polystyrene molecules under their natural configurations is briefly described as follows. A drop of dilute polystyrene ethyl bromide solution was dropped in hot water from a position a few inches above the water level. The solution droplet went into the hot water and was heated up before it came back to the water surface. As soon as the solution droplet came out of the water surface, it spread out radially and the polystyrene molecules formed a thin film on the water surface while the ethyl bromide evaporated away. The concentration of the polystyrene ethyl bromide solution varies depending upon the rate of evaporation of this solution. An approximate figure is a few percent of polystyrene by volume in ethyl bromide. The temperature of the hot water is not critical either. For polystyrene, around 80 degrees C, which is the second-order transition temperature of polystyrene, is desirable.

Under this kind of technical treatment, both uniform film and film as shown in Fig. 16 can be obtained. On the basis of this figure and the previous discussion, several generalizations concerning structure may now be attempted:

(a) Under certain conditions, polystyrene molecules can be grouped together as slender domains, (b) It is possible to orient these domains in one general direction, (c) Macroscopic fine structure of polystyrene domains may be observed.

However, there are still many interesting phases which are not understood and need further study. For instance, what is the relationship between the volumetric energy and the surface energy of each domain? Would these domains of polystyrene be present if mercury or other substance was used instead of water, as water molecules are of strong polar groups?

Now, let us consider the size of these rodlets or domains and how they compare with the length of the polymer chains. Figure 16 shows a section of the film magnified some 19,000 times so that a distance of 2 cm would correspond to about 10^{-4} cm or 1 micron. Most of the rods appear to be of the order of 1 micron in length and perhaps one-tenth of a micron in thickness. Now, in polystyrene the primary bonds are C−C bonds having a distance l between them of 1.54A and an angle α between the valence bonds of 109°. The maximum length of an extended chain is then given by[8]

$$L_{\max} = Nnl \sin(\alpha/2),$$

where N is the number of monomer units and n is the number of bonds per unit. For polystyrene, the extended

[8] E. Guth and H. M. James, Ind. Eng. Chem. 33, 624 (1941).

FIG. 15. Electron micrograph of polystyrene film formed by dropping polystyrene ethyl bromide solution on top of cold water surface showing large domain structure.

length of the polymer chain is of the order of 1.6×10^{3}A, or roughly $\frac{1}{6}$ of a micron. Thus, the length of the rodlets is approximately four to six times the length of the polymer chains. The width of the domains is, on the other hand, some 100 times larger than the approximate distance across adjacent chains (10A) and hence if the domain is interpreted as a large conglomeration of more or less parallel chains, it should contain several thousand or more depending on how closely they are packed.

In the polymerized solid state rather than in film form it may well be that the individual long-chain molecules of polystyrene are not arranged in this fashion although this type of structure has been suggested and the polymer chains treated as long rods.[9,10] Regardless

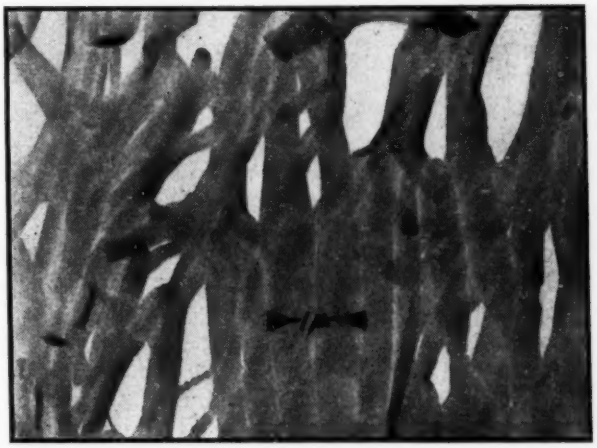


FIG. 16. Electron micrograph of polystyrene film formed by dropping polystyrene ethyl bromide solution in hot water showing oriented domain structures of polystyrene.

[9] H. Mark and G. S. Whitby (Editors), "Collected Papers of W. H. Carothers on High Polymeric Substances," *High Polymers* (Interscience Publishers, Inc., New York, 1940), Vol. I.

[10] K. H. Meyer, "Natural and Synthetic High Polymers," *High Polymers* (Interscience Publishers, Inc., New York, 1942), Vol. IV.

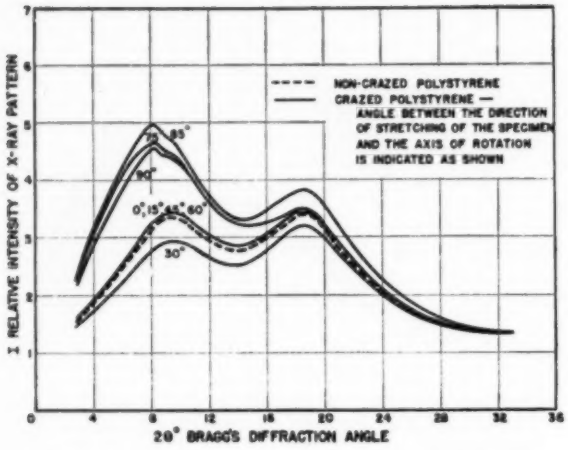

NON-CRAZED POLYSTYRENE
CRAZED POLYSTYRENE —
ANGLE BETWEEN THE DIRECTION OF STRETCHING OF THE SPECIMEN AND THE AXIS OF ROTATION IS INDICATED AS SHOWN
I RELATIVE INTENSITY OF X-RAY PATTERN
2θ° BRAGG'S DIFFRACTION ANGLE

FIG. 17. X-ray photometer curves of crazed polystyrene.

of this situation, Fig. 16 reveals that the natural configuration of polystyrene molecules, at least under certain environments, is one of more or less oriented, intertwined cylindrical domains. Before additional generalizations concerning structure are attempted, results of x-ray investigations of the structure of crazed and non-crazed material will be discussed.

X-RAY INVESTIGATION OF CRAZED AND NON-CRAZED POLYSTYRENE

Both crazed and non-crazed polystyrene samples have been subject to x-ray and electron diffraction examination. The diffraction picture[5] reveals that polystyrene is amorphous in structure. Very little information can be obtained from this picture except to note that the existence of the two somewhat diffuse rings is indicative of the existence of two spacings in some fairly definite regular pattern. Since only one of these rings (the outer one) appears in x-ray pictures of the monomer the other is characteristic of the polymer and may represent either chain to chain spacing or else link to link spacing.

If, however, the x-ray spectrometer is used to obtain intensity measurements, the results indicate a variation of the intensity of the diffracted rings and a shift of the peak of the inner ring or polymerization ring (see Fig. 17). If the peaks indicate some characteristic spacings of the C—C chains, the variation of these spacings with angle is shown in Fig. 18. This shift of the peaks and the variation of the intensity of the diffracted ring for different directions of the axis of rotation relative to the direction of stretching indicate the existence of orientation and "crystallization" of the material.[5,11,12] It should also be noted that no such indications of orientation are found if the non-crazed portions of a specimen are examined by means of the x-ray diffraction.[5]

[11] P. H. Hermans and A. Weidinger, J. Polymer Sci. **IV**, 135 (1949).
[12] P. H. Hermans and A. Weidinger, J. Polymer Sci. **IV**, 709 (1949).

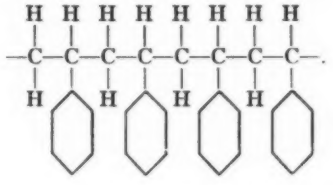
CRAZING AND MOLECULAR STRUCTURE

In the preceding sections, the results of light microscope, electron microscope and x-ray examinations of crazed and non-crazed polystyrene material, as well as of some naturally oriented polymeric film, have been presented. The studies were made on polystyrene or polymethyl methacrylate specimens which had been permitted to craze under the action of stress and time. The crazed material is found to consist of very fine, very smooth more or less parallel openings lying in planes at right angles to the principal tensile stress direction. For a given magnitude and time duration of stressing, the size of the openings and their density varies with the particular polymer under consideration. The openings are finer and much more dense for polystyrene than for methyl methacrylate. Crazing as such is not confined to these two materials, but probably occurs in all linear polymers. It is visible, however, only in transparent materials.

The studies also show that there are differences in orientation between the crazed and non-crazed portions of a linear polymer, although these differences are slight when the crazing is produced by long-time loading at room temperature conditions. There seems also to be some evidence supporting the view previously presented[5] based on observations of behavior of linear polymers under various types of mechanical testing, that polystyrene molecules, at least under certain environments, may group together in more or less cylindrical domain form, with the various long-chain molecules essentially parallel to one another within the domain. The molecular forces across these groups of chains are probably not very large and the precise configuration will depend on which structure gives the minimum free energy. The polymer chains are held together by secondary van der Waals forces and possibly by some side group interlocking.

EFFECT OF DOMAIN STRUCTURE ON CRAZING

Many of the facts pertaining to the inception, growth, and location of crazing and crazing cracks in polystyrene seem to be explainable in terms of orientation and the separation of domains of polystyrene molecules under the action of tensile stress. Consider first the accepted chemical structure of the polystyrene molecule. It is a series of styrene monomers arranged regularly in the following manner



Because of the presence of the large benzene rings around the main C—C chains, it is difficult for these chains to slide in the longitudinal direction of the chains.

However, it should be relatively easy to separate parallel chains by pulling them in their perpendicular directions. It may also be possible under the action of stress for one domain to separate and form several shorter domains if the original domain is many lengths longer than the average length of the long-chain molecules.

The styrene polymer presumably possesses no unused valence bonds except possibly at chain ends. Thus, there are essentially no primary bond linkages across polymer chains and hence the only forces across chains should be relatively weak non-valence or van der Waals London type forces.[13] These forces may be sufficiently large to maintain a rod-like structure when no external forces are acting but under applied transverse tensile stress the polymer chains within the rods would tend to separate from one another. Other evidence that the polymer of styrene is of a thread-like structure arises from consideration of its Raman spectra.[14]

When material having a randomly directed domain type of structure is placed under tensile stress, the various domains tend to orient themselves according to the state of stress applied. Those which are perpendicular to the direction of applied stress will theoretically not be able to orient and in virtue of their weak transverse holding power will open under the action of applied stress. Along the external surfaces of the material these openings will grow more easily than those in the inside, both because of stress concentrations arising from the existence of surface flaws and also because the restrictive force fields of neighboring domains of molecules will probably be somewhat less for domains located on the open surface. When these openings become sufficiently large (many times the dimension of the wave-length of light), they are seen as macrocracks by the naked eye. However, long before the openings become large enough to be individually seen, and when still in the form of microcracks or submicroscopic openings, they may still have the ability to reflect light.[15] In other words, these openings may be sufficiently numerous and sufficiently regular in their general orientation pattern to cause a large proportion of the light striking them to be reflected. Crazing is then evident in the form of surface blushing rather than visible cracks.

When a rigid linear polymer is subjected to tensile stress at room temperature, the tendency of the original randomly directed domains to orient is greatly restricted by the internal molecular interactions of other groups and chains. If, however, the applied stress is high and is maintained for some time, the domains originally

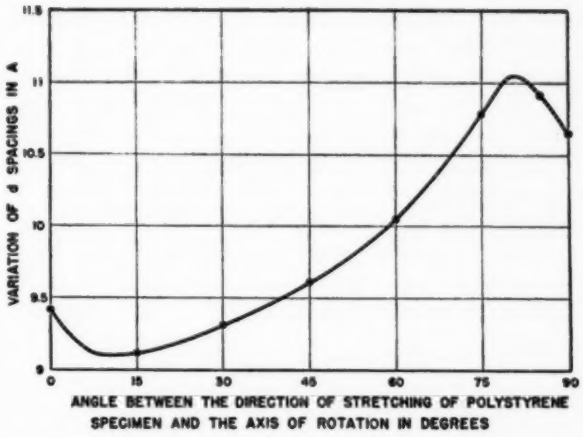


FIG. 18. Variation of polymerization ring of crazed polystyrene.

directed at right angles to the direction of stressing will open up and will initiate the phenomenon of crazing. It is usually possible for many crazing cracks to occur before fracture because the micro-openings soon run up against other domains or groups of molecules not so favorably oriented for separation and then they are automatically halted. A somewhat similar view of the development and growth of crazing cracks has been presented by Maxwell and Rahm.[4]

The mechanism of crazing is now becoming more apparent. Domains or groups of molecules which are at right angles to the applied stress open up first but do not immediately produce failure or a fracture crack as we might expect from prior knowledge of the behavior of brittle materials because they soon run into domains of regions not so oriented. On the submicroscopic or even microscopic scale, the material is by no means homogeneous and the crazing planes are stopped by irregularities in the orientation pattern very much like slip lines in metals are stopped by crystal boundaries lying between differently oriented crystallites. The flatness of the crazing surface (see Fig. 9a) suggests a slow propagation of crazing caused mainly by the joining together of the many microscopic openings induced by separation of the properly oriented domains of the material in their longitudinal directions. The difference in elevation of crazing planes, noted to be about one-fifth of a micron in Fig. 12, is of the same order of magnitude as the extended length of a polymer chain and serves as an indication of the distance the microcrack can extend without running into differently oriented domains.

EFFECT OF ORIENTATION ON CRAZING AND STRENGTH

The effects of orientation on crazing and strength are intimately tied in with the effects of molecular structure. When groups of molecules tend to join together to form domains, it is the opening up of those groups oriented more or less at right angles to the direction of stressing that initiates the development of microcracks or crazing.

[13] A. E. Van Arkel, *Molecules and Crystals* (translated by J. C. Swallow) (Butterworths Scientific Publications, London, and Interscience Publishers, Inc., New York, 1949), Chapter IX.

[14] H. Mark and R. Raff, "High Polymer Reactions, Their Theory and Practice," *High Polymers* (Interscience Publishers, Inc., New York, 1941), Vol. III, p. 73.

[15] S. W. Hawkins and R. B. Richards, J. Polymer Sci. IV, No. 4, 515 (1949).

Also, since the domains can carry considerable stress in their longitudinal direction but only very little stress in the transverse direction, it is clear that the orientation of the domains will greatly affect the strength properties of the polymer.

In general, two kinds of orientation concern us, differing greatly in degree. One is orientation produced by hot stretching and the other by cold stretching. Hot stretching is usually carried out above the second-order transition temperature and only a very small force is required to stretch the material to thousandths of percent of its original length. Evidently, above the second-order transition temperature, thermal agitation energy is great enough to overcome the secondary attractive forces which hold the various polymer chains in more or less tight formation. These chains are then free to slide past one another as well as to orient in the direction of the applied tensile stress. It is thus to be expected that exceedingly large percentages of orientation (in terms of original length) can be produced when the stretching is done at these high temperatures. If the previous picture of the development of crazing is correct, it is also to be expected that when tensile stress is applied in the direction of orientation to specimens previously oriented to some great degree (~1000 percent), crazing, as such, should be absent. This is so because the specimen will no longer contain domains oriented at right angles to the tensile direction and hence no separation of domains is to be expected. Fracture, when it occurs, should be of the fiber type.**

For cold stretching at ordinary room temperature, only little orientation occurs prior to fracture. In polystyrene specimens, failure occurs at a strain of approximately 1.45 percent in./in. and hence only local short range orientation can be accomplished. Somewhat greater orientation might be expected from relatively slow creep tests but even here in terms of percent increase over original length about 1.5 percent seems to be the maximum obtained.

Evidently, the restrictive force fields of neighboring molecules are too great at room temperatures to allow much orientation. That which does occur is also not homogeneous over the entire cross-sectional area of the specimen. On the outside regions of the specimen where surface flaws and the growth of microcrazing cracks affect the stress distribution, x-ray evidence indicates that more orientation takes place than in the central non-crazed portions of the specimen.

The effect of orientation of groups of molecules on strength is a problem which should be susceptible to statistical treatment. It might also be used to check the concept that the polymer molecules are usually grouped together in the form of cylindrical domains bonded together by long-range intermolecular forces with the

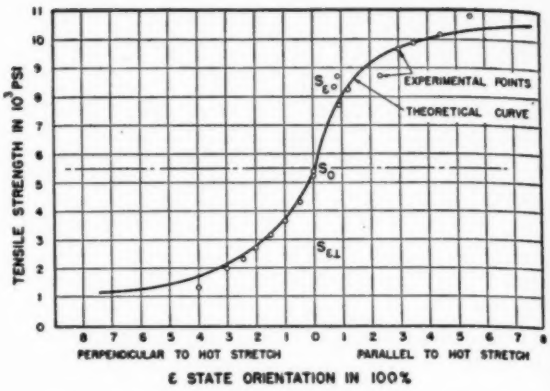


FIG. 19. Effect of uniaxial orientation on ultimate strength of polystyrene.

long chain molecules more or less parallel in each individual domain. For an amorphous polystyrene specimen, these domains of the polymer molecules are randomly oriented. The ultimate strength of this material tested at 77°F±2°F, 50 percent R.H. with the rate of straining 20×10^{-6} sec.$^{-1}$, has been determined experimentally to be approximately 5400 p.s.i. The ultimate strength of oriented material can be shown mathematically to be[16]

$$S_\epsilon = 2\frac{(1+\epsilon)^{\frac{1}{2}}}{1+(1+\epsilon)^{\frac{1}{2}}}S_0, \tag{1}$$

where S_0 is the strength of this material without any preferred orientation, ϵ is the measure of the unit strain in the direction of orientation, and S_ϵ is the ultimate strength of this material oriented at a unit strain change of ϵ in the direction of orientation. When $\epsilon\to0$, S_ϵ is seen to approach S_0 as $(1+\epsilon)^{\frac{1}{2}}/[1+(1+\epsilon)^{\frac{1}{2}}]$ approaches $\frac{1}{2}$. It can also be shown that for the material with perfect orientation in one direction, i.e. $\epsilon=\infty$, S_ϵ approaches $2S_0$ since the expression

$$\left[\frac{(1+\epsilon)^{\frac{1}{2}}}{1+(1+\epsilon)^{\frac{1}{2}}}\right]_{\epsilon\to\infty} = 1.$$

This condition as well as Eq. (1) in general seems to be in satisfactory agreement with experimental studies of oriented polystyrene made by Bailey[17] and hence seems to lend further support to the domain concept of molecular structure.

If oriented material is tested in tension in the direction perpendicular to the direction of orientation, the statistical treatment is somewhat more complicated. The expression found for the ultimate strength is[16]

$$S_{\epsilon\perp} = 2\frac{(1+\epsilon)^{\frac{1}{2}}}{(1+\epsilon)^3-1}\cdot\frac{2}{\pi}[K-E]S_0, \tag{2}$$

** These observations are in accord with the work of Bailey (see reference 17). Dr. Dietz informs us that similar observations have been experimentally noted in tests on highly oriented polystyrene specimens made at the M.I.T. Plastics Laboratory.

[16] C. C. Hsiao, "Effect of Uniaxial Orientation on Strength and Crazing of Some Linear High Polymers." (It is hoped to submit this paper for publication elsewhere in the near future.)

[17] J. Bailey, India Rubber World 118, 225 (1948).

where K and E are the values of Legendre's complete integrals of the first and second kind respectively. The value $S_{\epsilon\perp}$ varies from S_0 to 0 as ϵ varies from 0 to ∞.

Both Eqs. (1) and (2) for the strength parallel and perpendicular to the stretching direction are plotted in Fig. 19 for various values of percent stretch. It is seen that the predicted results and the experimental results previously reported by Bailey[17] are in good agreement. Small discrepancies that exist between the theoretical and experimental results may be due to differences in rates of straining applied to the specimen.[18]

The general principle employed in the development of the mathematical theory of the effect of orientation on the strength properties of polystyrene is that a domain can contribute its full strength in essentially only one direction, i.e., the direction of its length. Thus, the strength of a linear high polymer is not uniformly distributed from the microscopic point of view. The strength of the material depends upon the orientation of the domains. When stress is applied, some of the unfavorable oriented domains will open up while the more favorably oriented domains will carry the load. The micro-openings gradually grow and join together particularly if the stress is applied for a long period of time. For reasons mentioned previously, they usually start first on the surface and then propagate from the surface of the material toward the center. When they become sufficiently large and numerous enough to noticeably reflect visible light, crazing is said to have occurred.

CONCLUSIONS

A considerable amount of evidence is now available to draw the general conclusion that crazing is an actual mechanical separation of polymer chains or groups of chains under the action of tensile stress. The stress system need not be external as crazing can be produced by internal stresses caused by shrinkage, absorption, radiation, or combinations thereof. Even when the proper stress conditions are present, however, the separations may be too small to be seen either by the naked eye or by a high powered light microscope. The phenomenon of crazing is probably common to many solids other than transparent linear high polymers but, if the material is not transparent, crazing is not observed.

Crazing usually starts on the surface in a region where the polymer chains are oriented at right angles to the direction of applied tensile stress and then gradually penetrates inward. Crazing will not occur if the material is fully oriented in the direction of stretching and it can be reduced for randomly oriented materials by means of suitable surface coatings. Crazing, or even visible crazing cracks, will not necessarily produce fracture but will reduce the total load the specimen can carry. Compression tends to close existing crazing openings and to partially heal them but cannot restore the material to its former state.

For polystyrene, many of the facts pertaining to crazing are explicable in terms of orientation of molecular domains containing many polymer chains held together in more or less parallel fashion by secondary forces. The same type of structure seems to provide a satisfactory model for calculating the effect of orientation on the strength properties of the material.

ACKNOWLEDGMENTS

This project is a joint undertaking with the Materials Branch of the ONR and has been carried on under their sponsorship. The authors wish to acknowledge with many thanks the advice, counsel, and assistance which they received from Dr. W. F. Bartoe of Rohm and Haas Company, and from Dr. T. F. Bates, Mrs. M. V. Black, Dr. K. E. Russell, Dr. W. J. Harrington, Miss E. J. Richeson, and Miss C. Y. Yuan of The Pennsylvania State College.

[18] C. C. Hsiao and J. A. Sauer, "Effect of Strain Rate on the Tensile and Compressive Stress-Strain Properties of Polystyrene," Technical Report No. III, Contract No. N6onr-269, Task Order VI, for Mechanics and Materials Branch, ONR (January, 1950).

End-Cooling of Power Tube Filaments*

JOHN W. CLARK** AND RALPH E. NEUBERT†
Research Division, Collins Radio Company, Cedar Rapids, Iowa
(Received February 24, 1950)

The differential equation defining the relation between temperature and distance in that portion of a vacuum tube filament which is cooled by both conduction and radiation is formulated. This equation has been solved by numerical integration; the results are presented as a set of universal curves in a form convenient for use in vacuum tube design. The calculations agree well with some experimental measurements which were made to verify the theory.

INTRODUCTION

IN conventional power tube design the end cooling of the filaments by their supports is not particularly important and is usually treated by simple approximate methods.[1] This is due to the fact that the filaments in such tubes are very long compared to their diameter, so the power lost by conduction is much smaller than that radiated. Power tubes intended for use at very high frequency must have their filaments not too long as compared with the shortest wave-length at which they will be operated; accordingly as the radio industry continues to demand more power at high frequency the tube designer is forced to use correspondingly short filaments in which conduction cooling cannot be neglected.

This paper presents a method for designing the end portion of a filament so that the temperature distribution in that portion of the filament affected by conduction can be readily determined. The solutions are presented as a family of universal curves for convenience in design work. While this method is intended primarily for use in connection with the design of short filaments, it may be used for filaments of any length.

FORMULATION OF THE PROBLEM

In order to fix the ideas, let us first consider the simple case of a filament of cross section area A, perimeter P, and resistivity ρ which is cooled by radiation only. Then in any incremental length of filament the power lost by radiation must equal the Joule heat developed in that same increment of length by the current passing through it. That is

$$I^2R=I^2(\rho dx/A)=P\gamma dx, \tag{1}$$

where I is the current passing through the filament and γ is the radiation coefficient, i.e. the number of watts per square centimeter radiated by the filament. The equation is set up in terms of P, the perimeter, so it will apply regardless of the shape of the wire. It is convenient to write this equation in terms of the current density $J\equiv I/A$, and of a shape factor $F\equiv P/A$. In terms of these new variables,

$$J^2\rho=F\gamma. \tag{2}$$

This simple equation determines the operating temperature of a radiation-limited filament. Ordinarily it is required to determine the current to produce a given temperature in a wire of given cross section. Since both ρ and γ are known as functions of temperature for any material which might be used to construct a filament, it is a straightforward matter to solve the problem.

Let us now consider that portion of the wire which is so close to the support that it is not at uniform temperature. This portion of the wire will lose heat both by radiation and by conduction. Figure 1 defines the geometry of the problem. The heat carried across a boundary at x must be the difference between the Joule heat generated and the heat lost by radiation between x and 0. The point $x=0$ is defined as the value of x where $dT/dx=0$. Beyond this point, radiation alone suffices to carry away the I^2R heat, as formulated in Eq. (2). This energy balance can be formulated in the following equation, letting K be the coefficient of thermal conductivity:

$$\int_0^x \frac{I^2\rho(T)}{A}dx-\int_0^x P\gamma(T)dx=AK(T)\frac{dT}{dx}. \tag{3}$$

Rewriting (3) in terms of current density and shape factor, we have the fundamental integro-differential equation of temperature distribution in a filament cooled by both conduction and radiation:

$$\int_0^x J^2\rho(T)dx=\int_0^x F\gamma(T)dx+K(T)\frac{dT}{dx}. \tag{4}$$

Now $\rho(T)$, $\gamma(T)$, and $K(T)$ are all known functions of temperature. Accordingly we can differentiate (4) to get it in the form of a second order differential equation:

$$J^2\rho(T)=F\gamma(T)+K(T)\frac{d^2T}{dx^2}+\frac{dK(T)}{dx}\frac{dT}{dx}. \tag{5}$$

The remainder of this paper is concerned with obtaining solutions of Eq. (5).

* This work was supported in part by the United States Air Force through a subcontract between Collins Radio Company and the RAND Corporation.
** Now at Varian Associates, San Carlos, California.
† Now at Sylvania Electric Products, Inc., Emporium, Pennsylvania.
[1] L. R. Koller, *The Physics of Electron Tubes* (McGraw-Hill, Book Company, Inc., New York, 1937), pp. 89–91.

SOLUTION OF THE TEMPERATURE EQUATION

The formal solution of Eq. (5) is

$$x(T)=\int_{T}^{T_H}\frac{KdT}{\left(\int_{T}^{T_H}2K(J^2\rho-F\gamma)dT\right)^{\frac{1}{2}}}. \qquad (6)$$

The solution of Eq. (5) is outlined in Appendix A.

Since ρ, γ, and K are all known functions of temperature, it is in principle possible to evaluate the integrals in (6) for any case of interest. This is a very laborious process and hence it is useful to calculate a set of generalized curves which can readily be used in any particular case.

The only metal for which detailed data are available for values of electrical resistivity, thermal conductivity, and specific radiation over the temperature range of interest is tungsten. Fortunately this is the metal most likely to be used for a power tube filament.

Equation (6) has been solved by numerical integration using the known values of the physical properties of tungsten. The numerical integration was performed by the Computations Group of the RAND Corporation, Santa Monica, California. This assistance is gratefully acknowledged.

The method used in the numerical integration is given in Appendix B. The results were computed with the aid of an auxiliary function K^* defined by the following equation:

$$x(t)=(10^4/J)K^*(t, t_H, \alpha), \qquad (7)$$

where t is temperature measured in kilo-degrees centigrade and the parameter α is the shape factor F divided by the square of the current density in amperes per square centimeter. That is

$$\alpha\equiv 10^6F/J^2. \qquad (8)$$

These calculations enable one to predict the temperature distribution in a tungsten filament in the range of interest in typical power tubes. Typical curves are shown in Fig. 2 for one specific case. The agreement between calculation and experiment is quite gratifying.

One is frequently interested in the distance from the cold filament support to the point where radiation cooling becomes controlling, rather than in the detailed variation of temperature up to this point. This point has previously been identified as the point where $T=T_H$; it is here that dT/dx vanishes and the useful part of the filament (as far as electron emission is concerned) starts. Let us denote the value of x corresponding to $T=T_H$ by l. Then l is the length of that part of the filament in which conduction cooling causes a variation in temperature.

It frequently happens that l is fixed by other design considerations. Since l is an implicit function of t_H, F, and J, it is possible to use the calculated values to determine values of F and J consistent with the given

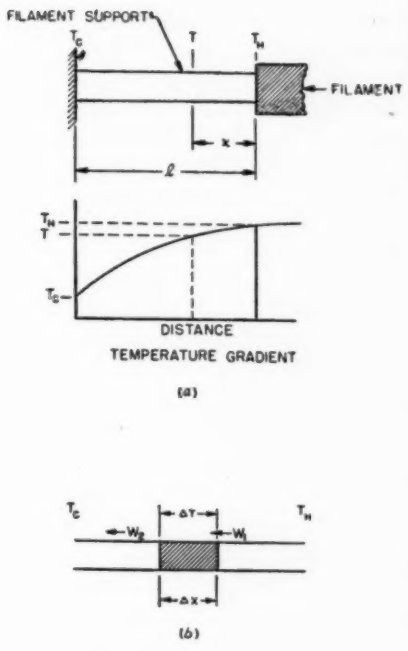


FIG. 1. The geometry of the problem.

values of t_H and l. To facilitate this process, the curves of Fig. 3 were prepared. An inspection of Eq. (7) shows why this single family of curves can give all the required information. Remembering that $l\equiv x(t_H)$, we obtain from (7):

$$l=(10^4/J)K^*(t_H, \alpha). \qquad (9)$$

Thus it is possible to plot one family of curves with lJ as ordinate, with α as abscissa, and with T_H as a parameter. It is convenient to use $(lJ)^{-1}$ as ordinate rather than lJ; this has been done in Fig. 3. On this plot the heavy lines are fixed values of T_H, the temperature of the emitting portion of the filament. The abscissa is $\alpha\equiv 10^6F/J^2=10^6PA/I^2$. The ordinate is $10^4(Jl)^{-1}=10^4A/Il$. The light lines are lines of constant $100Pl/I$.

One can enter the plot with given values of any four of the five quantities T_H, P, A, I and l and obtain the corresponding value of the fifth. Usually other considerations fix T_H, I, and l; and if the shape of the wire is fixed, P and A are of course not independent. This means that it may be necessary to make either the shape or the area of that portion of the filament where temperature variations exist different from that of the filament proper. This can be regarded as a procedure for controlling the position of the beginning of the emitting region; the curves of Fig. 3 give one all the information necessary to do this.

DETERMINATION OF PHYSICAL CONSTANTS

The method outlined above can also be used to determine the three physical constants ρ, K, and γ of metals at elevated temperatures. The writers have not done this, but feel that a brief outline of the procedure would be appropriate. In view of the paucity of data concerning the physical properties of metals at high

temperatures, it would be very desirable for someone to undertake these measurements.

A test piece (of circular cross section) at least ten diameters long would be mounted between electrodes of accurately controlled temperature and of accurately known spacing. Data would be obtained of brightness temperature and current as a function of applied voltage. A new sample of the same wire would then be mounted in the same way, with the electrodes moved so that the sample is now only about four diameters long; it must still have a short radiation-limited area at its center at the highest operating temperature. The same data as before are taken. It is now easy to obtain resistance *vs.* temperature curves for each sample, and by subtraction to obtain resistance *vs.* temperature for the mid-portion of the filament which was cooled by radiation only. Thus we can calculate ρ as a function of temperature, and by the use of Eq. (2) we can calculate the specific radiance γ.

If we are dealing with a metal for which the relation between brightness temperature and blackbody temperature is not known, this will have to be determined in a separate experiment.

The determination of K, the thermal conductivity, can be accomplished with the same set-up. We now use a still shorter wire, so that radiation limiting does not occur. We then measure the mid-temperature, and the distance Δx from the midpoint of the wire to a point where the temperature is ΔT degrees less than the mid-temperature. Writing Eq. (4) as a difference equation, and solving for K, we find

$$K=(J^2\rho-F\gamma)(\Delta x)^2/2\Delta T. \tag{10}$$

The factor 2 in the denominator occurs because we are averaging K over the interval in question.

APPENDIX A

To solve the second-order non-linear differential equation

$$J^2\rho(T)=F\gamma(T)+K(T)\frac{d^2T}{dx^2}+\frac{dK(T)}{dx}\frac{dT}{dx}, \tag{5}$$

where $\rho(T)$, $\gamma(T)$, and $K(T)$ are known tabulated functions.

Rearranging,

$$\frac{d^2T}{dx^2}+\frac{K'(T)}{K(T)}\left(\frac{dT}{dx}\right)^2=\frac{J^2\rho(T)-F\gamma(T)}{K(T)}. \tag{11}$$

We have written K' for dK/dT and have made use of the fact that $dK/dx=dK/dT\cdot dT/dx$. Now let $dT/dx=u$; then $d^2T/dx^2=udu/dt$, and our equation becomes

$$u\frac{du}{dT}+\frac{K'(T)}{K(T)}u^2=\frac{J^2\rho(T)-F\gamma(T)}{K(T)}. \tag{12}$$

Let $m=u^2$; then $udu/dt=\frac{1}{2}dm/dt$, and

$$\frac{dm}{dt}+2\frac{K'(T)}{K(T)}m=2\frac{J^2\rho(T)-F\gamma(T)}{K(T)}. \tag{13}$$

This is a linear differential equation of the first order. The solution is:[2]

$$m[K(T)]^2=2\int_T^{T_H}k(T)[J^2\rho(T)-F\gamma(T)]dT. \tag{14}$$

Returning to our original variables,

$$\left(\frac{dT}{dx}\right)^2=\frac{2}{[K(T)]^2}\int_T^{T_H}K(T)[J^2\rho(T)-F\gamma(T)]dT. \tag{15}$$

Equation (15) can be formally solved for x, giving the integral

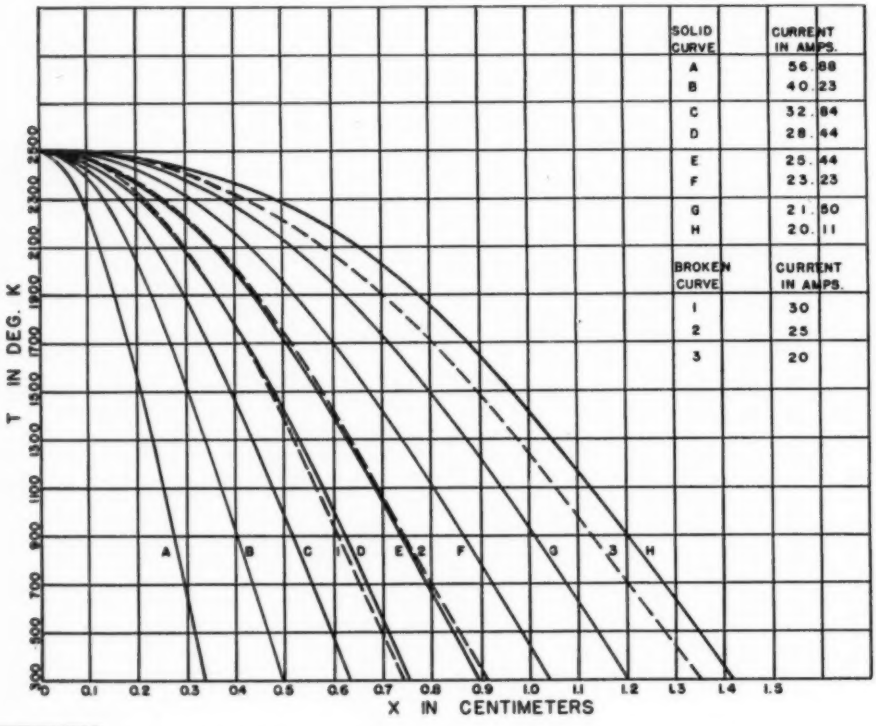


FIG. 2. Comparison between measured and calculated values of temperature distribution in a tungsten filament 0.020 in. diameter. The solid curves are calculated, while the dashed curves are measured values.

[2] F. S. Woods, *Advanced Calculus* (Ginn and Company, Boston, 1932), p. 220.

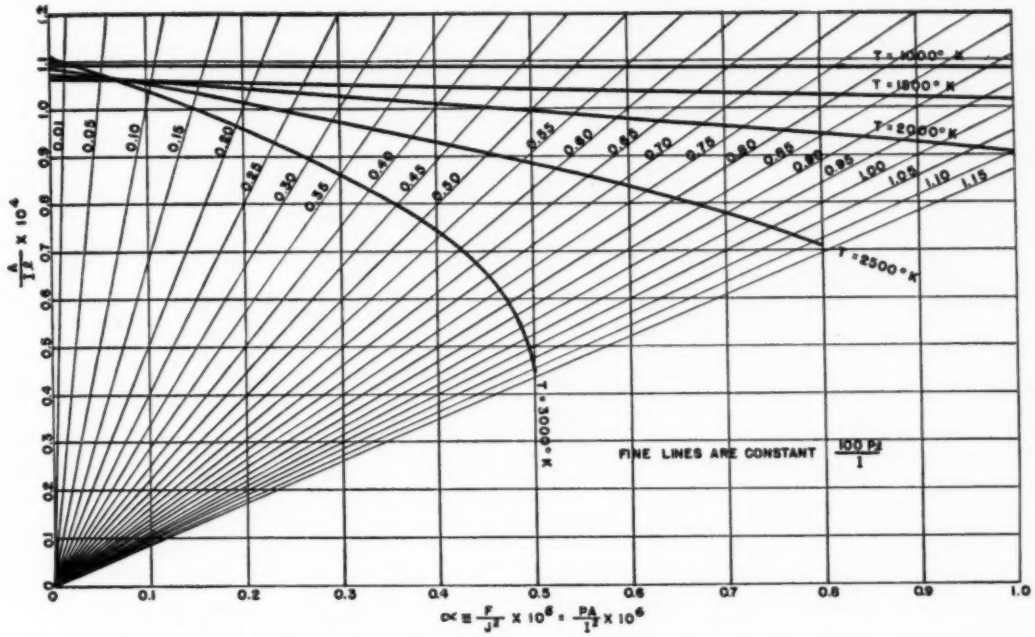


FIG. 3. Universal curves for determining the length of the conduction cooled portion of a tungsten filament.

solution

$$x(T)=\int_{T}^{T_H}\frac{K(T)dT}{\left(\int_{T}^{T_H}2K(T)[J^2\rho(T)-F\gamma(T)]dT\right)^{\frac{1}{2}}}. \tag{6}$$

APPENDIX B

The numerical integration of Eq. (6) was carried out as follows:

The physical constants of tungsten as functions of temperature can be approximated by the following equations:

$$K(T)=1.346-1.6\times10^{-4}T, \tag{16}$$

$$\gamma(T)=1.818\times10^{-19}(\log T)^{35.3}, \tag{17}$$

$$\rho(T)=1.51\times10^{-6}\left(\frac{T}{100}\right)^{1.21}. \tag{18}$$

The data from which these empirical equations were derived was obtained from Jones and Langmuir[3] for electrical conductivity and radiation constants, and from Coolidge[4] and Osborn[5] for thermal conductivity.

The analytic functions are introduced into the integrand of (6) and the integration performed by machine. In order to present the results compactly, two new variables are introduced:

$$t=T/1000, \tag{19}$$

$$\alpha=10^6F/J^2. \tag{20}$$

It is apparent from the form of (6) that it can be written

$$x(t)=(10^4/J)K^*(t,t_H,\alpha); \tag{7}$$

Eqs. (6) and (7) together constitute the definition of the function K^*. This function was computed with the aid of the approximate analytical formulations given in (16), (17) and (18) for $T_H=1000°K$, $1500°K$, $2000°K$, $2500°K$, and $3000°K$, for appropriate fixed values of α. The results were used in plotting the design curves given in Fig. 3.

[3] H. A. Jones and I. Langmuir, "The characteristics of tungsten filaments as functions of temperature," Gen. Elec. Rev. **30**, 312–313 (June, 1927).

[4] *Handbook of Chemistry and Physics*, 30th edition (Chemical Rubber Publishing Company, Cleveland, Ohio, 1947), p. 1865. Converted to watts/cm²/°C from Coolidge's data at 18°C.

[5] R. H. Osborn, J. Opt. Soc. Am. **31**, 428–431 (June, 1941). There appears to be a genuine discrepancy between the room temperature value of K and that at elevated temperatures. The empirical Eq. (16) fits Osborne's data and does not pass through the point given at room temperature. Further work is needed to resolve this discrepancy.

Space-Charge Effects in Electron Beams and Their Reduction by Positive Ion Trapping*

E. G. Linder and K. G. Hernqvist**
RCA Laboratories Division, Radio Corporation of America, Princeton, New Jersey
(Received April 19, 1950)

The spread of electron beams caused by space-charge forces is discussed. Under certain conditions ions formed from the residual gas may be trapped in the beam. An equilibrium condition may be established in which the electron and ion densities are equal, and then space-charge forces will be neutralized. Under these conditions high current densities may be produced at low voltages. A theory of ion trapping is discussed, and the equilibrium condition is formulated.

Experimental data are presented which were obtained by the application of microsecond pulses to the beam. This technique is described, and its advantages and possibilities are mentioned. Data are given on beam spread as a function of current, voltage and pressure. Data on the improvement due to ion trapping are included. An increase of current density by a factor of 30 was observed with the structures tested. Results are included on ion build-up time as a function of pressure, and on beam noise and stability in the presence of trapped ions.

I. INTRODUCTION

SPACE charge is the chief limiting factor in the production of high current-density electron beams. On account of space charge, repulsive forces exist so that beams parallel at their origin, become divergent, or beams originally convergent fail to converge. This effect appears to have been studied first by E. E. Watson[1] who, in 1927, derived the expression for the divergence of an initially parallel beam. Numerous other workers[2–8] have further developed Watson's results, and some have devised methods of controlling the spread. In 1940 Thompson and Headrick[9] extended the theory to the case of convergent beams. Field, Spangenberg, and Helm[10] showed in 1947 that a universal curve could be drawn giving beam spread in terms of certain general parameters.

This article covers the following topics: (a) Theory of space-charge spread, (b) Ion trapping and space-charge neutralization, (c) Pulse technique and apparatus, (d) Experimental data, (e) Beam stability.

M.k.s. units are used throughout, except when otherwise specified. The abbreviation SCN will be used throughout for "space charge neutralized," or "space charge neutralization."

II. THEORY OF SPACE-CHARGE SPREAD

Thompson and Headrick[9] have derived an expression for the spread of a convergent beam giving the shape, or profile, for systems of circular transverse symmetry. They found that the radius of the beam at any point is

$$r = R_0\epsilon^{-\pi\epsilon_0\frac{m}{e}\frac{v}{I}(v_{ro}^2-v_r^2)},$$
$$= R_0\epsilon^{-S(V_{ro}-V_r)/V}, \quad (1)$$

where R_0 = initial beam radius, I = beam current, ϵ = natural logarithmic base, ϵ_0 = permittivity of free space, m = electron mass, e = electron charge, v = beam velocity, v_r = radial velocity, v_{ro} = initial radial velocity, V = beam voltage, V_r = radial velocity in equivalent volts, V_{ro} = initial radial velocity in equivalent volts, and $S = (8\pi^2\epsilon_0^2e/m)^{\frac{1}{3}}V^{\frac{1}{2}}/I$. The minimum radius is obtained by putting $V_r = 0$, that is,

$$r_m = R_0\epsilon^{-SV_{ro}/V}. \quad (2)$$

The longitudinal (axial) position of any surface electron is found to be

$$z = 2S^{\frac{1}{2}}R_0\epsilon^{-SV_{ro}/V}\int_{(SV_r/V)^{\frac{1}{2}}}^{(SV_{ro}/V)^{\frac{1}{2}}}\epsilon^{x^2}dx, \quad (3)$$

and the axial position of the point of minimum radius is

$$z_m = 2S^{\frac{1}{2}}R_0\epsilon^{-SV_{ro}/V}\int_0^{(SV_{ro}/V)^{\frac{1}{2}}}\epsilon^{x^2}dx. \quad (4)$$

A typical set of curves computed from Eqs. (1) and (3) is given in Fig. 1. Profiles b are plotted for a series of beam currents from 0 to 60 ma. The beam voltage is 300, the source radius $R_0 = 0.3$ cm, and the focal distance with no spreading is $z_0 = 0.8$ cm. The curve cc is the locus of the points of minimum radius.

In the same figure, a partition A is shown which contains a circular aperture a, concentric with the beam. By use of the beam profiles it is evident that the amount of current passing through this aperture may be computed. Thus it is seen that for currents up to about 9.0 ma, the total amount traverses the aperture, whereas at 20 ma less than 25 percent passes through, and at 60 ma only about 3.5 percent.

* Presented at the Symposium of the Division of Electron Physics, Cambridge, Mass., June 18, 1949.
** Scandinavian-American trainee.
[1] E. E. Watson, Phil. Mag. **3**, 849 (1927).
[2] M. Knoll and E. Ruska, Ann. d. Physik **12**, 607 (1932).
[3] F. D. Fowler and G. E. Gibson, Phys. Rev. **46**, 1075 (1934).
[4] A. Bouwers, Physica **2**, 145 (1935).
[5] B. von Borries and J. Dosse, Arch. f. Electroteck. **32**, 221 (1938).
[6] A. V. Haeff, Proc. I.R.E. **27**, 586 (1939).
[7] L. P. Smith and P. L. Hartman, J. App. Phys. **11**, 220 (1940).
[8] J. R. Pierce, J. App. Phys. **11**, 548 (1940).
[9] B. J. Thompson and L. B. Headrick, Proc. I.R.E. **28**, 318 (1940).
[10] L. M. Field, K. Spangenberg, and R. Helm, Elec. Communication **24**, 108 (1947).

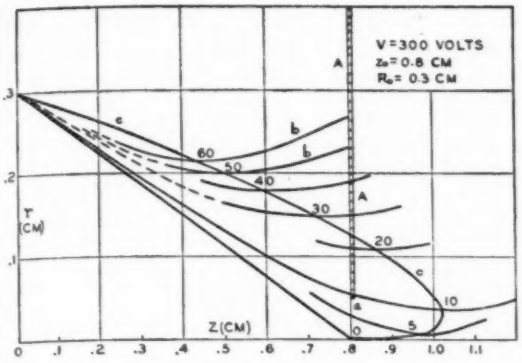


FIG. 1. Typical computed beam profiles.

If I_a is the current traversing the aperture, and I is the total beam current, the beam efficiency may be defined as I_a/I. This is determined by the ratio of the aperture area to the beam cross-sectional area, i.e.,

$$\mathcal{E}=I_a/I=r_a^2/r_1^2,$$

where r_a is the aperture radius and r_1 is the beam radius at the aperture.

A set of curves showing beam efficiency, thus computed, is given in Fig. 2, for beam voltages of 200, 300, and 400 volts. These curves, as well as those of Fig. 1, are idealized in that they are based on the original Thompson and Headrick assumptions, and furthermore no account has been taken of any aberrations, or effects of initial velocities.

By inspection of Eqs. (2) and (4) it is seen that r_m and z_m may be written as functions of the quantity $V^{\frac{1}{2}}/I$, since, by assumption, V_{ro} is proportional to V. It is not evident by inspection that Eqs. (1) and (3) also may be written as functions of this same parameter. However it has been shown by Field, Spangenberg, and Helm[10] that universal beam profiles may be drawn in terms of this parameter. Hence curves such as those of Fig. 1, may be considered universal, in so far as one may be drawn for each value of $V^{\frac{1}{2}}/I$, for any given geometry.

Thus, the curves of Fig. 2, also may be combined into a single universal curve by plotting $I/V^{\frac{1}{2}}$ along the horizontal axis instead of I. Such a curve is shown in

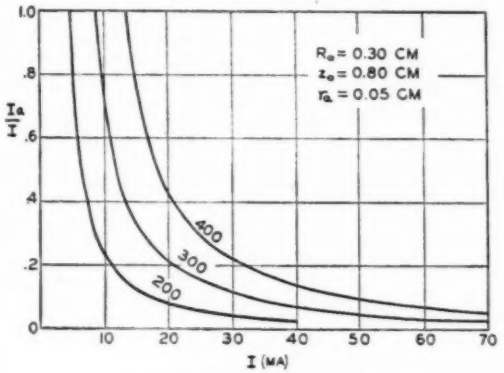


FIG. 2. Theoretical beam efficiency.

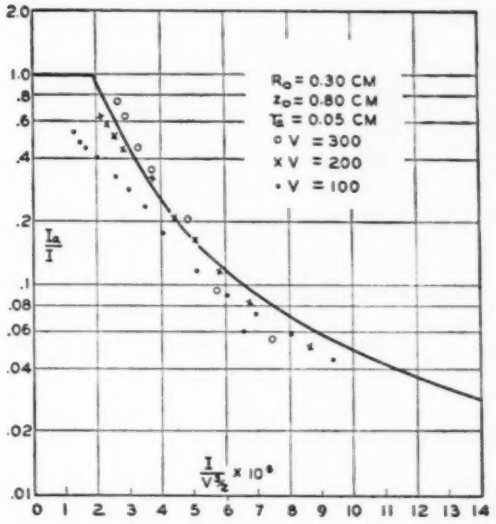


FIG. 3. Universal beam efficiency curve with experimental points at three different voltages.

Fig. 3. The experimental data shown in this figure will be discussed later.

III. ION TRAPPING AND SPACE CHARGE NEUTRALIZATION

In this section the prevention of space charge spread will be discussed. The method used here is based on the neutralization of electron space charge by positive ions. The first similar application of positive ions appears to have been made by J. B. Johnson[11] in 1922 in connection with a cathode-ray oscillograph tube in which beam spread was reduced by admitting gas into the tube. However, the conditions differed widely from those under discussion in this article, and did not involve the trapping of ions as described hereinafter.

A more recent application by O. Heil[12] has been described but many details are not available. The most detailed study was made by Field, Spangenberg, and Helm.[13] Here ions formed by collisions between beam electrons and residual gas molecules were trapped in long, high voltage beams. It was also shown that the number of residual gas molecules at pressures used in ordinary high vacuum tubes was sufficient to supply ions for beams of high current density. However, again, conditions and methods differed considerably from those of the present work.

An understanding of the ion-trapping process may be had with the aid of Fig. 4, which shows the potential variation across the tube diameter. At the top, a cross section of a cylindrical tube and beam is illustrated, where D is the cylindrical wall, and b is the beam, see

[11] J. B. Johnson, J. Opt. Soc. Am. 6, 701 (1922).
[12] O. Heil, *German War Time Research and Development in Klystrons*, U. S. Dept. of Commerce, Publication Board Report No. 52,348.
[13] See reference 10. This same work was printed privately in 1942. See report on "The Production and Control of Electron Beams," Electrical Engineering Dept., Stanford University.

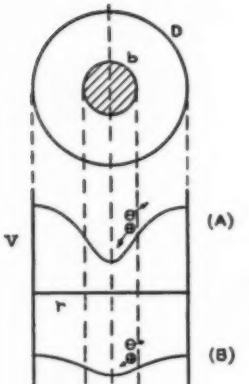


FIG. 4. Potential distribution across beam with various degrees of SCN.

Fig. 6. Below at (A) is shown a typical potential variation for an unneutralized beam. In the case of a cylindrical, parallel beam the variation is approximately parabolic within the beam, and logarithmic between the beam and the wall, (see Smith and Hartman[7]). Thus there is a potential minimum at the beam center.

This depression extends along the longitudinal axis, but decreases as the grid or aperture are approached, and becomes negligibly small at these electrodes. Thus there is formed a potential valley surrounded by higher regions on all sides.

Such a potential valley acts as an ion trap, as shown in Fig. 4(A). Positive ions, formed by collisions between beam electrons and residual gas molecules, move toward the beam center, and are trapped. On the other hand, low velocity electrons, formed by the ionization process, are forced away from the center of the beam.

The resulting accumulation of ions in the beam causes a progressive neutralization of electron charge, and consequently, the potential rises and the depression decreases, as shown at (B). If, eventually, the positive ion density equals the electron density, the potential becomes uniform, as at (C), then ion-trapping obviously ceases, and an equilibrium condition is established. Actually, under most circumstances, there will remain a slight potential depression of the order of a few tenths of a volt just sufficient to overcome the thermal energies of the ions. This will be discussed in detail below. Furthermore, electrons formed by ionization will continue to be lost from the beam, even with no potential depression, since they will encounter no resisting forces.

When the electron space charge is thus neutralized, the electron paths become straight and radial. The total current then passes through the aperture, except insofar as there are limitations due to unconsidered factors, such as structural aberrations and initial velocities.

The rate of formation of ions in an electron beam is given by

$$(dN_2/dt)_1 = \pi r^2 p P(V) n_1 v_1, \tag{5}$$

where p is the gas pressure, P is the number of electrons (or ions) formed per unit length, per electron, at unit pressure, n_1 is the electron density, N_2 is the number of ions per unit beam length, and v_1 is the electron velocity.

The rate of loss of ions will be computed by assuming that the trapped ions form an ion gas. This seems plausible in view of the long lifetime of the ions in the trap and the consequent possibility of the attainment of an equilibrium state. The rate of loss will then depend upon the ion gas density at the wall and the average normal velocity component there. The density will be controlled by the Boltzmann law and at the wall it will be

$$n_2 = n_2' \epsilon^{\Delta V e/kT_2},$$

where n_2' is the ion density at the beam center, T_2 is the gas temperature, and ΔV is the potential difference between the beam center and surrounding cylindrical wall. In the case of an electron beam, the potential is depressed, and ΔV is such as to make $\Delta V e/kT_2$ negative for positive ions. The average normal velocity at the wall is, from kinetic theory,

$$v_2 = \left(\frac{V_{02} e}{2\pi M}\right)^{\frac{1}{2}}.$$

Thus the rate of ion loss may be written

$$\left(\frac{dN_2}{dt}\right)_2 = 2\pi R n_2' \epsilon^{\Delta V/V_{02}} \left(\frac{V_{02} e}{2\pi M}\right)^{\frac{1}{2}}. \tag{6}$$

The net rate of change of number of ions may therefore be written

$$\frac{dN_2}{dt} = \left(\frac{dN_2}{dt}\right)_1 - \left(\frac{dN_2}{dt}\right)_2 = \pi r^2 p P n_1 v_1 - 2\pi R n_2' \epsilon^{\Delta V/V_{02}} \left(\frac{V_{02} e}{2\pi M}\right)^{\frac{1}{2}}. \tag{7}$$

Throughout the trapping period the first term remains constant, however, at the beginning of the process, n_2' is small, also $|\Delta V| \gg V_{02}$, hence the second term is small. As the trapped ion density n_2' increases, $|\Delta V|$ also decreases, and the second term increases. At equilibrium the two terms must be equal, then $dN_2/dt = 0$, and Eq. (7) becomes,

$$\pi r^2 p P n_1 \left(\frac{2Ve}{m}\right)^{\frac{1}{2}} = 2\pi R n_2' \epsilon^{\Delta V/V_{02}} \left(\frac{V_{02} e}{2\pi M}\right)^{\frac{1}{2}}.$$

Cancelling terms, and solving for p, gives

$$p = \frac{R n_2'}{\pi^{\frac{1}{2}} r^2 P n_1} \left(\frac{m}{M} \frac{V_{02}}{V}\right)^{\frac{1}{2}} \epsilon^{\Delta V/V_{02}}. \tag{8}$$

This represents the equilibrium condition for a potential depression in the beam. It is evident that ΔV is a function of pressure, and that at a certain pressure, $\Delta V=0$. Under this condition also $n_2'=n_1$, and then

$$p_{(\Delta V=0)}=\frac{R}{\pi^{\frac{1}{2}}r^2P}\left(\frac{m}{M}\frac{V_{02}}{V}\right)^{\frac{1}{2}}. \tag{9}$$

Only at this pressure is there accurate space charge neutralization; at other pressures n_2' is not exactly equal to n_1. However, under practical conditions, such as are of interest here, ΔV is usually less than a volt, and then $n_2'\approx n_1$ for beam currents of several milliamperes or more. This arises from the fact that the potential depression is caused by the unneutralized charge n_1-n_2', and since these are both large numbers, n_1-n_2' may be large even though n_1/n_2' is nearly unity.

When the pressure exceeds the value given by Eq. (9), ΔV changes sign and the potential depression becomes a potential increase. This is due to the formation of ions at a rate greater than that at which they can escape without an assisting electric field.

In this case ions are no longer trapped, but electrons are, and thus it is now possible to set up equations for electrons, similar to Eqs. (5) and (6), representing the rates of electron formation and loss.

The rate of flow of electrons to the wall will be given by

$$\left(\frac{dN_1}{dt}\right)_1=2\pi Rn_1'\epsilon^{-\Delta V/V_{01}}\left(\frac{V_{01}e}{2\pi m}\right)^{\frac{1}{2}}. \tag{10}$$

The rate of formation of electrons in the beam will be

$$\left(\frac{dN_1}{dt}\right)_2=\pi r^2pPn_1\left(\frac{2Ve}{m}\right)^{\frac{1}{2}}. \tag{11}$$

Here n_1' is the density of the electron gas at the beam center, that is, the density of the trapped electrons. This does not include the beam electrons, which are not trapped, do not form a gas, and do not obey the Boltzmann law. Also V_{01} is the electron gas temperature in equivalent volts.

Equating Eqs. (10) and (11) and solving for p yields,

$$p=\frac{Rn_1'}{\pi^{\frac{1}{2}}r^2Pn_1}\left(\frac{V_{01}}{V}\right)^{\frac{1}{2}}\epsilon^{-\Delta V/V_{01}}. \tag{12}$$

This equation is valid only for potential rises in the beam, whereas Eq. (8) is valid only for potential depressions. For $\Delta V=0$ they should both be valid. Equating them, and putting $\Delta V=0$, yields,

$$\frac{n_1'}{n_2'}=\left(\frac{m}{M}\frac{V_{02}}{V_{01}}\right)^{\frac{1}{2}}. \tag{13}$$

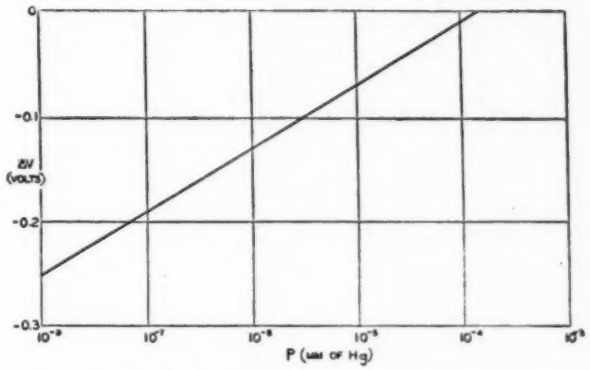


FIG. 5. Variation of potential depression with pressure.

If this is written

$$n_1'\left(\frac{V_{01}}{m}\right)^{\frac{1}{2}}=n_2'\left(\frac{V_{02}}{M}\right)^{\frac{1}{2}},$$

it is seen that this means that the electron and ion currents escaping from the beam, due solely to thermal motion, are equal, when $\Delta V=0$.

Equation (12) can not now be evaluated for other values of ΔV since n_1' is an unknown function of ΔV. Further work must be done to determine this function.

However, Eq. (8) for the potential depression case, can be evaluated, assuming $n_1=n_2'$. This is plotted in Fig. 5, for the following conditions: $R=0.5$ cm, $r=0.1$ cm, $P=7.7$, $M/m=52{,}000$, $V=300$ volts, and $V_{02}=0.026$ volt. These values correspond approximately to a set of experimental conditions to be discussed later. The values of R and r are averaged for a conical beam. P and M/m are taken for nitrogen. The ion gas temperature is assumed to be 300°K. These data give

$$p=1.5\times10^{-4}\epsilon^{\Delta V/0.026}.$$

It is seen from this that as the vacuum improves, the equilibrium potential depression increases in order to decrease the rate of ion escape from the trap. However for all practical vacuums the depression is very small compared to the beam voltage.

The build-up time of the ion density and the aperture current I_a may be computed from Eq. (7). However, this computation is complicated by the fact that the variation of ΔV during build-up is not known. Fortunately the term which contains ΔV does not become important until the final stages of build-up, when ions begin to escape from the trap. It is then responsible for the tail of the curves, as shown in Figs. 9 and 11. An approximation may be made by neglecting this term and using only the first term of Eq. (7). The values thus obtained will be too small, but will be sufficiently accurate for many purposes. If this is done, there results,

$$dn_2/dt=pPn_1v_1$$

since

$$N_2=\pi r^2n_2.$$

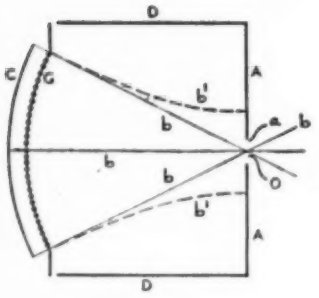


FIG. 6. Schematic diagram of ion-trapping gun.

The ion build-up time τ is the interval required for n_2 to equal n_1, i.e.,

$$\tau=\int_0^{\tau} dt=\int_0^{n_1}\frac{dn_2}{pP(V)n_1v_1}.$$

But $v_1=(2Ve/m)^{\frac{1}{2}}$, and also, if the number of ionized gas molecules is small compared to the total number of residual gas molecules in the beam-region, so that p can be considered constant during the ionization build-up period, it follows that

$$\tau=\frac{(m/2e)^{\frac{1}{2}}}{pP(V)n_1V^{\frac{1}{2}}}\int_0^{n_1}dn_2,$$

or

$$\tau=\frac{(m/2e)^{\frac{1}{2}}}{pP(V)V^{\frac{1}{2}}}.$$

If τ is measured in microseconds, p in mm of mercury, and $P(V)$ in number of ions formed per cm, per electron, at a pressure of 1 mm of mercury, this becomes

$$\tau=\frac{0.0169}{pP(V)V^{\frac{1}{2}}}. \qquad (14)$$

Thus τ varies inversely as the pressure, is independent of current, and is a function of voltage which can be specified only when $P(V)$ is known.

In the above treatment it is assumed that the number of residual molecules in the beam region is large compared to the number of ions required, that is, to the number of electrons in the beam. This condition is usually easily consistent with electron beam tube operating conditions.

IV. PULSE TECHNIQUE AND APPARATUS

The construction of the tubes used in the tests is shown schematically in Fig. 6. The cathode C comprises a segment of a spherical surface with its center of curvature at 0. Immediately in front of the cathode is the grid G, which also is a spherical segment having its center of curvature at 0. The aperture a, through which it is desired to pass current, is placed at 0. The disk A, cylinder D, and grid G form a shielded enclosure, which is field-free when G, A, and D are at the same potential. Under this condition, and in the absence of space charge, electrons will follow radial trajectories, as shown by lines b. But in the presence of space charge, divergence will occur, as illustrated by dashed lines b'.

A scale drawing of an actual tube is shown in Fig. 7. Parts are lettered to correspond to similar parts of Fig. 6. In addition there is shown here the focusing electrodes B, and the Faraday cage F to collect the aperture current I_a. The tests were made with a tube of demountable design in which the position of the electron gun and the Faraday cage could be adjusted by external knobs, working through metal bellows. This adjustment could be either lateral or longitudinal.

The greater part of the data obtained in this study was obtained by use of a pulse technique. It is believed that this is the first application of pulse methods to the investigation of electron beams. It has proved to be very superior to d.c. methods, eliminating several sources of error and yielding additional information. These advantages will be discussed below.

The circuit used is illustrated schematically in Fig. 8. The tube elements are cathode C, grid G, aperture disk A, and Faraday cage F. The pulser produced a square pulse of width T_2 variable from $\frac{1}{2}$ to 5 microseconds, and of spacing T_1 variable from 50 to 5000 microseconds. The pulse amplitude was adjustable up to 400 volts.

A typical oscillogram is sketched in Fig. 9, where grid voltage (above) and aperture current (below) are plotted against time. From a to b all electrodes have the same potential with respect to the cathode, and the current is constant. At b, a pulse is applied to the grid of sufficient amplitude to bring its potential down to zero. This stops the cathode emission current, and it also applies an ion-sweeping field between the grid and

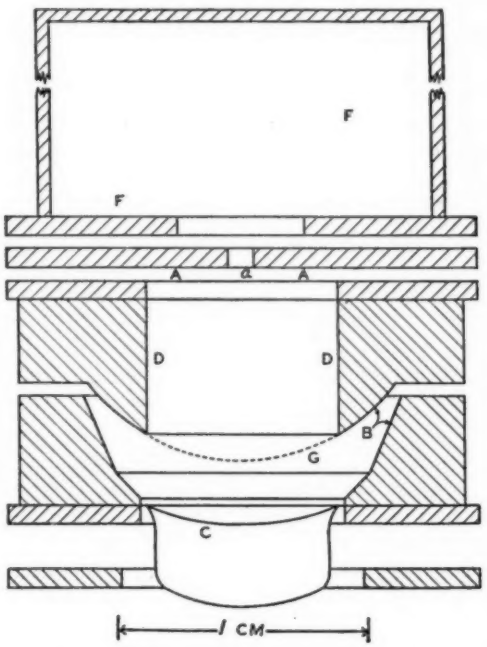


FIG. 7. Scale drawing of ion-trapping gun.

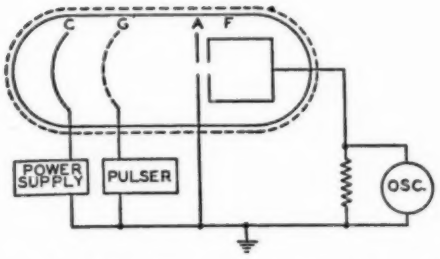


FIG. 8. Schematic diagram of tube and circuit.

the aperture for the duration of the pulse from c to d. At d the pulse is terminated and the grid-potential and the cathode emission current return to their initial values. The aperture current however rises only to e, which is less than the initial value. The reason for this is that the positive ions were swept out of the beam region during the application of the pulse, consequently at its termination, and the reformation of the beam, there is no space-charge neutralization. The current at e represents that passing through the aperture with the beam in a diverged state due to space charge. However, at point e ions again begin to be formed and trapped, hence there is an increase in current during the time interval from e to f. At the latter point the current has attained its original value, which represents that obtained with complete space-charge neutralization.

This method has several advantages over a dc method in which ions are swept out by applying a steady sweeping voltage on the grid. (1) Electrons maintain constant velocity between grid and aperture, as is required for comparison with existing theories. (2) There is no defocusing, or other distortion of the beam, by nonuniform fields in the beam region. (3) There are no complications caused by secondary electron currents. (4) An ion-sweeping field as large as desired may be used. (5) Data on ion build-up time are obtainable.

V. EXPERIMENTAL DATA

1. Space Charge Spread

The decrease in aperture current due to space charge spread may be obtained directly from oscillograms such as that of Fig. 9. The ratio of the current without SCN to that with, is the ratio of lengths ed/bc. The numerical value of the current at point b may be measured with a meter while the pulser is switched off. All other desired currents may likewise be metered except the current corresponding to point e. This is obtained from the ratio of ed/bc.

Measurements of this sort were made to determine the beam efficiency, $\mathcal{E}=I_a/I$, as a function of beam current I, gas pressure p, and beam voltage V. Typical oscillograms are shown in Figs. 10–12.

The variation of I_a/I with beam current shows directly the effect of space charge in causing beam spread. Oscillograms for a series of beam currents are shown in Fig. 10. Here the beam voltage and pressure are substantially constant. As the current is varied, point e (see Fig. 9) changes, moving upward with decreasing current. This shows that the space charge spread is smaller for smaller currents. Data taken from such oscillograms are plotted in Fig. 13, where beam efficiency, I_a/I is plotted against beam current for three different beam voltages. The rapid drop in efficiency due to beam spread is apparent.

As pointed out above, these data may be presented in the form of a universal curve by plotting against $I/V^{\frac{1}{2}}$, instead of against I alone. This has been done in Fig. 3. The solid line has been computed from the Thompson-Headrick theory; the points have been obtained from ed/bc ratios on oscillograms such as those of Fig. 10. These data at different voltages show a systematic drop in efficiency as the voltage decreases. This will be shown later to be due to the effect of initial thermal velocities. Both voltage and pressure effects will be discussed later.

2. Space-Charge Neutralization

In Fig. 13 are shown also the values of I_a/I for space charge neutralized beams. These are substantially horizontal lines, except for a small upward rise for small currents. This indicates that there is little or no beam spread as the current increases in this case. Over the range of the measurements there appears to be no limitation of current due to space-charge spread. The tests were not usually extended beyond $I=55$ ma at $V=300$, because of excessive heating of the grid, and danger of destroying it. However, for a period of a few seconds, an aperture current $I_a=68$ ma was obtained at 400 volts with a beam efficiency $E=86.8$ percent. The lower voltage tests terminate at values of current which represent the space charge limited values of current between cathode and grid. This is not a fundamental or serious limitation since it can be varied by changing the cathode-grid spacing. (There is no SCN effect in the cathode-grid region since it is not field-free and ion trapping does not occur.) It is seen from these data that the maximum gain obtained in these tests from SCN corresponds to an increase in I_a of about 30 times at $I=55$.

It also is evident from Fig. 11 that the gain due to

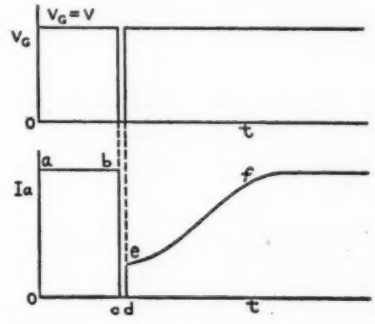


FIG. 9. Typical pulse shape showing applied grid voltage (above), and resulting aperture current (below).

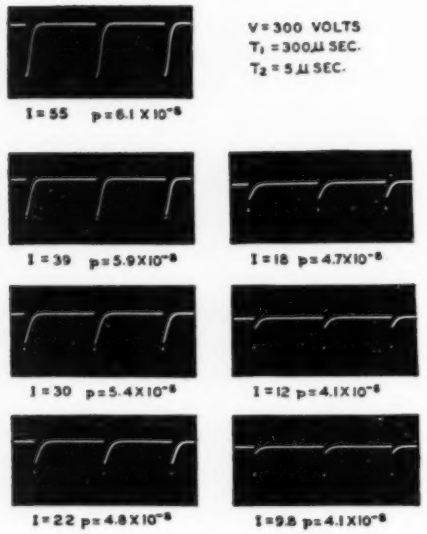


FIG. 10. Effect of current variation.

SCN is independent of pressure, since the ratio *bc/ed* does not vary with pressure. This is in agreement with the above theory.

3. Ion Build-Up Time and Pressure

The ion build-up time τ may be determined from oscillograms such as those of Fig. 9 by measuring the time difference between the points *e* and *f*. However, by using this method good agreement could not be expected when comparing the results with the simple theory used in deriving Eq. (14), because of some serious sources of error that cannot be overcome, namely: (1) The point *f* is very hard to localize on the scope picture because of the smooth transition to the equilibrium condition. (2) The tendency for the ions to escape from the beam region increases as the potential

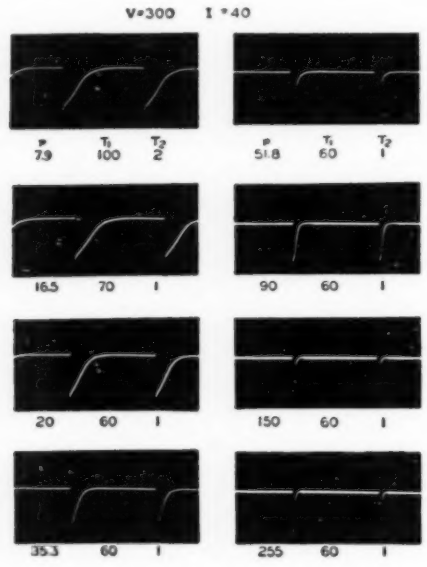


FIG. 11. Effect of pressure variation.

across the beam approaches the equilibrium state thus giving a longer ion build-up time than predicted by Eq. (14). These factors can be overcome by using another method, in which the build-up time is determined from the rate of formation of ions at the point *e* and from the knowledge of the number of ions required for full neutralization. The values thus obtained of course may differ somewhat from the practical ones, such as the time from *e* to *f* in Fig. 9, but they are consistent with the assumptions of Eq. (14) and are valid for comparison therewith. This method of obtaining τ has been used for all values given in this paper.

It is seen that, theoretically, according to Eq. (14), the build-up time varies inversely as pressure, inversely as $P(V)\ V$ and is independent of current. This latter independence can be seen to be substantially verified by the data of Fig. 10. No attempt at a more exact verification has been made.

The variation with pressure has been more completely checked. Pulse characteristics are shown in Fig. 11 for a range of pressures from 7.9×10^{-6} to 255×10^{-6} mm of mercury. These apply to a 40-milliampere, 300-volt beam. The decrease in build-up time is of an easily apparent magnitude in the photographs.

Values of τ obtained from such oscillograms are plotted against pressure in Fig. 14, on log-log paper. The solid lines are drawn with the theoretical slope of -1. The agreement lies within the experimental error. The 200-volt plot lies above the 300-volt one since the quantity $P(V)\ V$ of Eq. (14) is smaller in the former case. The values of $P(V)$ were taken for air.[14]

4. Variation of Beam Voltage

The build-up time is seen, from Eq. (14), to vary with beam voltage inversely as $P(V)\ V$. The function $P(V)$ in general has a value which rises to a maximum near a hundred volts and then drops slowly as V increases.[14] Multiplying by V shifts the maximum to a higher value. This function is inverted in the expression for τ, and results in a curve such as shown in Fig. 15, by the solid line. The experimental values, plotted as crosses, were obtained from oscillograms such as those of Fig. 12, by the method described in the preceding Section 3. Although the agreement is not quantitatively good, it is satisfactory in view of the assumptions involved. It is evident from known $P\ (V)$ functions that 300 volts represents an optimum voltage region, but that at higher voltages the build-up time would probably not be seriously larger.

At low beam voltages (below about 100 volts) it has been observed that the efficiency (I_a/I) drops very rapidly. This is illustrated in Fig. 16. It can be shown that this is caused by beam spread due to initial thermal velocities which become of importance when the beam velocity is small. To estimate this effect consider a

[14] See Engel and Steenbeck, *Elektrische Gasentladungen* (Berlin, 1932), p. 35.

small current, or an SCN case, so that space charge can be neglected. Assume a system such as shown in Fig. 17, where R_0 is the emitter radius, and r_2 the radius at the aperture, i.e., at the focal spot, due to initial velocities. Neglecting initial velocities, the focal spot would be a point, and the beam angle would be α, as shown.

For the system of Fig. 17, E. Ruska[15] has derived the expression

$$\frac{r_2}{R_0}=\frac{2\sin\beta}{\sin 2\alpha}\left(\frac{kT}{Ve}\right)^{\frac{1}{2}}.$$

To obtain the total spread put $\beta=\pi/2$. Also substitute the approximation,

$$\sin 2\alpha = 2R_0/z_0,$$

then

$$\frac{r_2}{R_0}=\frac{z_0}{R_0}\left(\frac{kT}{Ve}\right)^{\frac{1}{2}}, \quad \text{or} \quad \frac{r_2}{z_0}=\left(\frac{kT}{Ve}\right)^{\frac{1}{2}}. \tag{15}$$

Since the total current remains constant $\pi R_0^2 J_0 = \pi r_2^2 J$, or

$$J/J_0 = R_0^2/r_2^2, \tag{16}$$

where J_0 is the initial current density, and J the density at the focus.

From Eqs. (15) and (16),

$$J/J_0 = R_0^2/z_0^2(Ve/kT) = (Ve/kT)\tan^2\alpha.$$

For small α, $\tan\alpha \approx \sin\alpha$, so that this is in agreement with the expression of D. B. Langmuir.[16]

The fraction of current passing through an aperture of radius r_a is r_a^2/r_2^2, or from Eq. (15),

$$I_a/I = r_a^2/r_2^2 = r_a^2/z_0^2(Ve/kT). \tag{17}$$

This expression is valid, of course, only for $r_a \leq r_2$. Under this condition, and for a given geometry, the beam efficiency varies directly as V, the beam voltage.

This theoretical expression is plotted in Fig. 16 as a straight line through the origin. According to this simple theory the ratio I_a/I should rise linearly until it attains unity ($r_2=r_a$) after which it should remain constant. The departure of the experimental curve from the simple theory may be due in part to the neglect of the Maxwell-Boltzmann distribution in the theory. This distribution would cause a rounding off, instead of a sharp break, at $I_a/I=1$. Also there is spreading of the beam due to other factors, such as grid irregularities, which prevent the experimental value of I_a/I from reaching unity.

The slope of the line in the figure is drawn to equal that of the experimental curve at the origin and corresponds to a temperature of 1370°K. This is a little high, but the actual cathode temperature was not determined, and this value is close enough to indicate

[15] E. Ruska, Zeits. f. Phys. **83**, 684 (1933).
[16] D. B. Langmuir, Proc. I.R.E. **25**, 977 (1937).

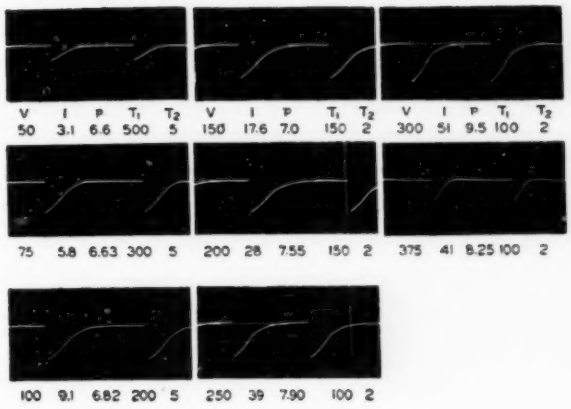


FIG. 12. Effect of voltage variation.

definitely that the drop in I_a/I is due to thermal velocities.

If the beam voltage is reduced below the ionization potential it is clear that the formation of ions will be affected and the neutralization of electron space charge may not be possible. Data taken in this region showed that as V increases from 9 to 10 volts there is an abrupt increase in I_a, corresponding, very likely, to the onset of ionization. However, as V is decreased from 10, the current drops smoothly until it reaches the original value at about 7 volts. Thus there is a hysteresis effect. This phenomenon is shown in the oscillogram of Fig. 18 wherein the horizontal scale was scanned by a 60 c.p.s. voltage, while the vertical represents I_a.

5. Comparative Results

Comparisons of SCN electron guns with several other types are given in Table I. The other types represent various previous attempts to obtain high current density, except, the 2K25 which is a widely-used commercial type, but not designed for particularly high density. The comparison is made on the basis of six quantities. As listed in order in the columns, they are: (1) total aperture current, (2) current density at the emitter, (3) current density at the aperture, (4) perve-

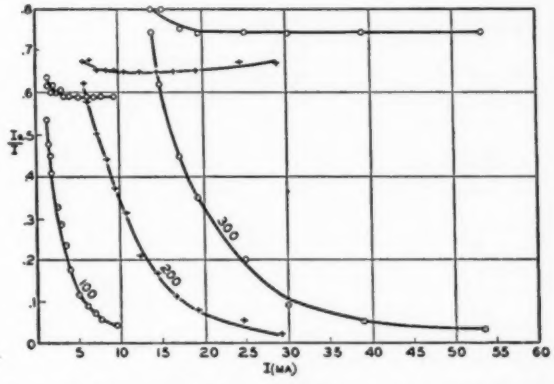


FIG. 13. Experimental beam efficiency curves with (above), and without ion trapping (below).

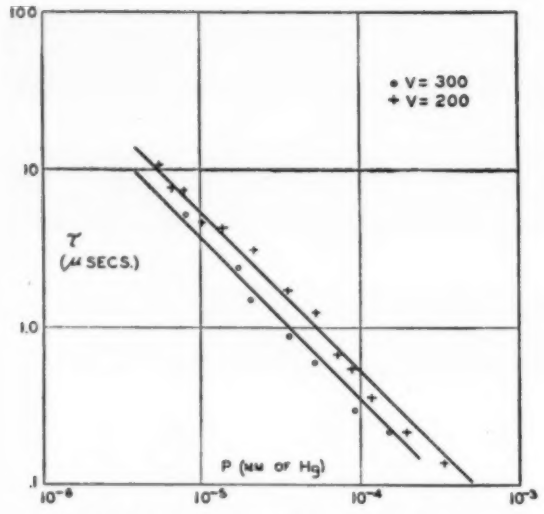


FIG. 14. Variation of ion build-up time with pressure.

ance$\times 10^6$, (5) concentration factor, or the ratio J/J_E, and (6) a figure of merit, F.

The figure of merit represents an attempt to indicate, in a manner independent of voltage and electrode size, the comparative ability of various guns to pass current through an aperture. To do this the Child-Langmuir law for parallel beams has been modified by introducing the factor F, i.e.,

$$I_a = 2.334\times 10^{-6}(V^{\frac{1}{2}}/x^2)a\cdot F,$$

where I_a is the current in amp./cm^2, V is the beam voltage, x is the distance from emitter to aperture in cm a is the aperture area in cm^2, and F is the figure of merit. For a parallel beam, $F=1$, and the greater the value of F the greater the aperture current compared to a parallel beam. Thus

$$F=\frac{I_a x^2 10^6}{2.334 V^{\frac{1}{2}} a}. \tag{18}$$

For space charge limited beams this is independent of

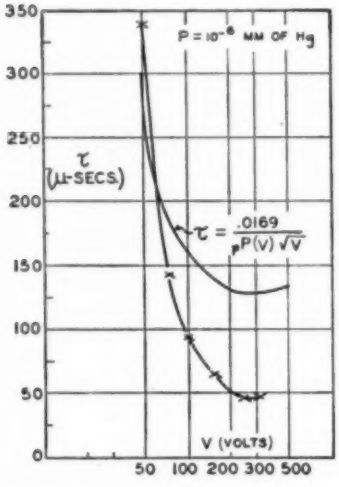


FIG. 15. Variation of ion build-up time with voltage, theoretical and experimental (X).

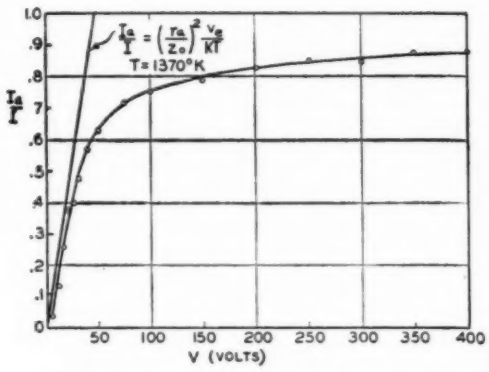


FIG. 16. Variation of beam efficiency with beam voltage. Current space charge limited between cathode and grid.

V, since I_a varies as $V^{\frac{1}{2}}$. Further, it is independent of gun size, since the ratio of x^2/a remains constant during scaling.

It is evident from the table that the SCN guns exceed all others in F value by a considerable factor. An important type of gun not listed in the table is one recently developed by O. Heil, for which very good results have been reported. This is not included, since published data are not yet available. It employs a very concave cathode, a wide-angle beam, and an optimized electrostatic focusing system.

VI. STABILITY

If reference is made to Fig. 11 it will be seen that at low pressures (10^{-5} mm of Hg) the oscilloscope trace is noise-free, but at higher pressures (90×10^{-6} and higher) there is a visible indication of noise. This is shown also in Fig. 19, which was taken with a beam voltage of 300, a current of 40 ma, and a pressure of 1.5×10^{-4} mm of Hg. It is noteworthy that the noise appears quite abruptly at a point where space charge neutralization apparently is completed. There is no evidence of it during the ion build-up period.

Reason to expect noise, and instability, under the observed conditions is found in the discussion of the equilibrium condition in Section III, and particularly in Eqs. (8) and (12). It was shown there that as the pressure increases, ΔV is at first negative, then passes through zero, and becomes positive. When ΔV is negative (potential depression) there is very nearly an exact equality of beam electron density and ion density. However, for ΔV positive, electron trapping starts,

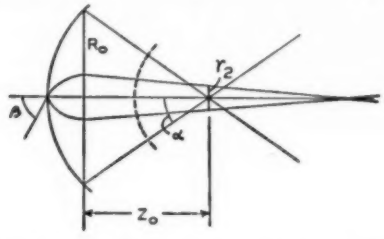


FIG. 17. Diagram showing beam spread due to thermal velocities in a structure with a focusing grid.

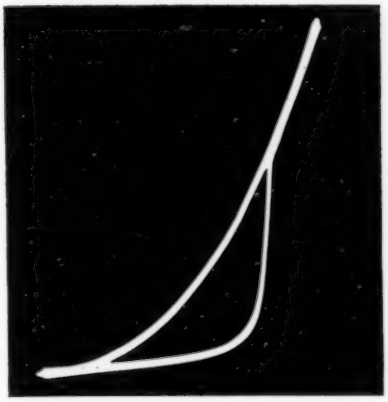


FIG. 18. Oscillogram showing variation of aperture current near ionization potential.

TABLE I. Comparative data on electron guns at 300 volts.

Type	I_a(ma)	$J_E\left[\frac{\text{ma}}{\text{cm}^2}\right]$	$J\left[\frac{\text{ma}}{\text{cm}^2}\right]$	Perveance $\times 10^6$	Conc. factor	F	Reference
Samuel	8.3	46.5	166	1.6	3.6	3.0	a
2K25	37.0	229.0	229	7.1	1.0	3.5	b
Lafferty	8.3	34.5	2200	1.7	64.0	38.0	c
SCN (1)	44.0	102.0	5700	8.5	56.0	470.0	
SCN (2)	49.0	113.0	6340	9.4	56.0	574.0	

[a] A. L. Samuel, Proc. I.R.E. **33**, 233 (1945).
[b] J. R. Pierce and W. G. Shepherd, Bell Syst. Tech. J. **26**, 460 (1947).
[c] J. N. Lafferty, thesis, University of Michigan, June, 1946.

and the electron and ion densities increase to some, at present, undetermined values. Furthermore, the temperature of the trapped electron gas in this case should be much higher than that of the trapped ion gas with ΔV negative. These conditions may well lead to noise production, and oscillations in the beam.

More specifically, noise has been found to be observable first at a pressure of approximately 10^{-4} mm of Hg. It rises to a maximum at 5×10^{-4}, and then remains constant to 10^{-3}, which is the upper limit of the experiments. It is seen that the threshold of 10^{-4} is in fair agreement with the theory, in that it is approximately at this value of p that ΔV changes from $-$ to $+$, as may be seen from Fig. 5.

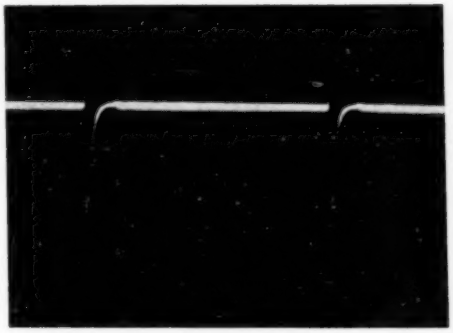


FIG. 19. Oscillogram showing appearance of noise in SCN beam.

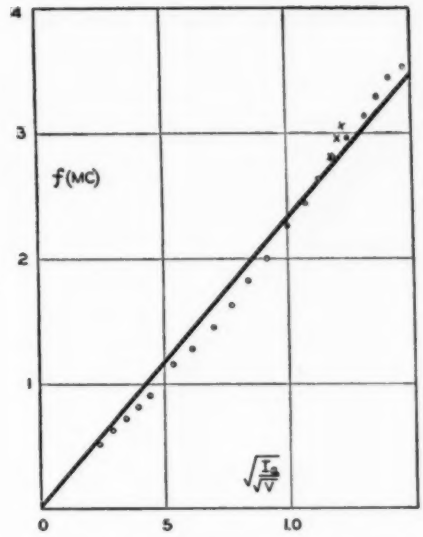


FIG. 20. Experimental beam oscillation frequency as a function of ion density.

Plasma ion oscillations also have been observed in the beam. These were found at the lowest pressures used, i.e., around 10^{-6}. They occurred up to $p=10^{-4}$ where they either ceased or were masked by the noise which begins at that pressure. It was not possible to determine which occurred.

For low pressures, the oscillation frequency was very distinct, and could be synchronized with an oscilloscope, or measured with a calibrated receiver. According to theory,[17] plasma ion oscillations have a frequency given by

$$f=\frac{1}{2\pi}\left(\frac{n_2e^2}{\epsilon_0 M}\right)^{\frac{1}{2}}, \tag{19}$$

where M is the ion mass and n_2 the ion density. A measure of the ion density could be obtained from the current density at the aperture, and the beam voltage. Thus

$$f=K(I_a/V^{\frac{1}{2}})^{\frac{1}{2}}, \tag{20}$$

where K is a constant depending on ion mass. The agreement with this formula is shown in Fig. 20. The straight line represents Eq. (20) for a value of molecular weight of 41. This result seems high because the value for N_2 is 28, and O_2, 32, but since the identity of the trapped ions is unknown no definite conclusion can be drawn. However, the agreement is sufficiently good, so that there is little doubt that plasma ion oscillations are occurring.

No plasma electron oscillations were detected although they were sought over the range of from 80 to 1000 Mc, with the same tube as exhibited the ion oscillations.

[17] I. Langmuir and K. T. Compton, Rev. Mod. Phys. **2**, 239 (1930).

Stress-Strain Properties of Natural Rubber under Biaxial Strain*

B. B. S. T. Boonstra
Delft, The Netherlands
(Received January 30, 1950)

Tube-shaped samples were stretched in a direction parallel to the tube axis and at the same time inflated to cause a tangential extension. In this way 600 percent axial and 300 percent tangential elongation could be attained simultaneously. The stresses were calculated from the pressure in the tube and the dynamometer readings.

The results are plotted as axial stress-strain curves at different tangential elongations ranging from 0–300 percent.

The curve for a higher tangential elongation lies at higher stresses than that for a lower tangential elongation; the differences are however larger than the theory predicts.

At 600 percent axial elongation stress relaxation was measured at different tangential elongations. It appears that the relaxation constant is independent of the tangential elongation.

INTRODUCTION

STRESS–STRAIN properties for rubber stretched in two directions have been reported by Ariano[1] and by Treloar.[2]

Both authors used rectangular test pieces and neither of them succeeded attaining high elongations because of the low tear resistance of rubber under biaxial stress.

High elongation was reached by the first author by inflating spherical samples, then however the strain in the majority of cases is about identical in the two directions. In this paper a system is described which allows stretching to high elongations in two directions independently of each other.

According to this method tube-shaped samples are stretched in the direction of the tube axis and at the same time inflated to cause a tangential elongation. The axial stress is partly measured on the dynamometer, the rest and the transverse stress is calculated from the air pressure in the tube. In this way stress-strain curves may be obtained for stresses parallel and at right angles to the tube axis as a function of the elongations in the two directions.

The theoretical equations extended to this case may then be checked. Besides, the decay of stress (axis-parallel) with time was measured and the influence of elongation perpendicular to this stress determined.

EXPERIMENTAL PROCEDURE AND RESULTS

Preparation of the Samples

Glass tubes of 16–25-mm diameter were repeatedly dipped in a latex dispersion of the following formula:

Rubber (in form of 40 percent latex)	100	
Colloidal sulfur	2	in form of 33 percent
Zinc oxide	2	dispersion with 0.2
Butyl zimate	1	percent Darvan
Latekoll (ammoniumpolyacrylate)	1	

After a layer of sufficient thickness was obtained (0.25–0.40 mm) the tubes were first dried at room temperature, then for 30 minutes at 80°C and finally vulcanized for 15–30 minutes at 110°C in live steam or hot air.

Dumb-bells punched out of these tubes showed the following properties:

Tensile strength	320–360 kg/cm²	
Elongation at break	750–800 percent	
Load at 300 percent elongation	17– 21 kg/cm²	
Hardness (shore *A*)	38	
Elasticity (shore *B*)	85	
Permanent set after 24 hours 200 percent elong.	measured after 1 hour	2.5 percent
	measured after 24 hour	0 percent

The rubber tubes are fastened on ribbed metal cylinders coated with a layer of vulcanized rubber (carbon black mix).

These cylinders can be attached to the dynamometer clamps; one of them has an air inlet connected to a manometer and a small hand operated air pump.

The tubes are elongated on a Schopper tensile tester (a) with air free to flow in and out, (b) inflated with air to the original diameter, (c) fastened on oversize clamp cylinders so that a constant transverse elongation can be maintained by pumping air into the tube.

The principle is illustrated in Fig. 1; some experimental values in Table I.

In calculating the stress in axis direction in case of definite transverse elongation, it has to be considered that the air pressure not only determines the tangential stress but also gives a contribution to the axial stress in form of pressure against the bottom and top plate of the cylinder.

Therefore the curves for inflated tubes do not begin at elongation zero but at a definite elongation determined by the air pressure necessary for the transverse elongation.

In Fig. 2 a number of curves, measured under different conditions is drawn, with axial stress, calculated on original cross section plotted against axial elongation.

The lowest curves are those measured without inflation, these practically coincide with the dumbbell curve;

* Communication No. 124. Rubber Foundation.
[1] R. Ariano, Rubber Chem. and Technol. **13**, 92 (1940).
[2] R. L. G. Treloar, Trans. Faraday Soc. **40**, 5 (1944); Proc. Phys. Soc. **60**, 135 (1948).

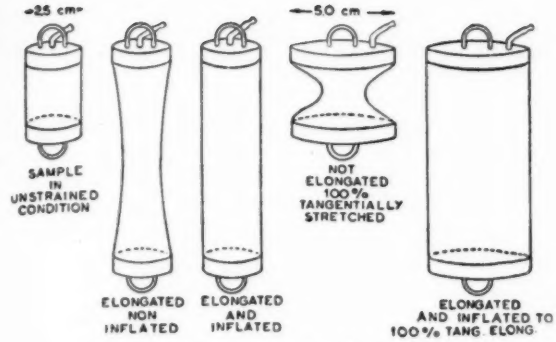


FIG. 1. Test samples for biaxial strain.

somewhat higher lies the curve for a tube which is not elongated sideways but inflated only to retain its original diameter. Considerably higher lie the curves for tubes with definite transverse elongations.

The pressure of the air, necessary to maintain a constant diameter of the tube at the different elongations is plotted as an average of several measurements against axial elongation in Fig. 3.

It will be observed in Fig. 3 that the pressure drops at the highest axial elongation for the 100 percent transverse elongation curve and also for the 300 percent elongation curve. Probably the rubber then has become so rigid that considerable changes in pressure are possible without much effect on the diameter of the tube.

The tangential stretching force on actual cross section is related to the pressure by the formula

$$\sigma = \frac{13.6 r \alpha_z}{1000(d/\alpha_s\alpha_z)} p \text{ kg/cm}^2 \text{ on actual cross section.} \quad (1)$$

Here r = original radius of the sample, mostly 12.5 mm, p = air pressure in cm Hg, α_z = axial extension ratio, α_s = transverse extension ratio, d = thickness of wall of the tube.

So at constant pressure and the same r, the actual stress in transverse direction is proportional to the

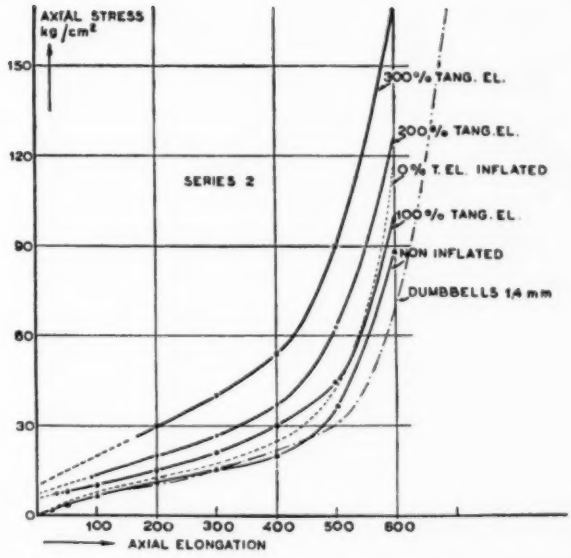


FIG. 2. Stress-strain curves at different tangential elongations.

square of the transverse elongation, (tr. el.)$^2 = \alpha_s^2$, this factor is four at 100 percent tr. el. and nine at 200 percent.

At the same pressures the actual stress at 200 percent tr. el. would be 9/4 of that at 100 percent tr. el. In reality however the stress ratio is lower, so it can happen that the pressure at 200 percent tr. el. is less than at 100 percent tr. el. (Fig. 3).

If the stress is plotted *vs.* the axial elongation, a set of curves is obtained as illustrated by Fig. 4.

In case the tangential tension is very small, for instance if the sample is not tangentially elongated, but only inflated to prevent contraction, the entity plotted in the figures was multiplied by ten; the stress in the direction of the axis is the actual value.

At 100 percent and higher transverse elongations the actual transverse stress has been plotted together with the stress in the direction of the axis, *vs.* axial elongation.

TABLE I. Some experimental values of tensile tests on inflated tube-samples (series 2) at 100 percent tangential elongation. Stresses are calculated in kg/cm² of the original cross section. r = original radius of the tube; α_z = Axial extension ratio; α_s = Tangential extension ratio; d = original thickness of the wall of the tube; p = air pressure in the tube in cm Hg.

Tangential elongation 100 percent		Axial elongation in percent 25	50	100	200	300	400	500	600
Average d = 0.033 cm	I Dynamometer axial stress kg/cm²	0.8	1.8	5.7	10.8	17.7	27.9	46.3	111.5
	II Pressure p in cm Hg	6.8	6.2	5.1	5.1	4.5	4.4	4.4	6.5
	III Axial stress calc. from $0.0272r/dp$ kg/cm² [a]	7.2	5.8	5.3	5.3	4.7	4.6	5.0	6.9
	I+III Total axial stress kg/cm²	8.0	8.6	11.0	16.1	22.5	32.5	51.3	118.4
	IV Tangential stress kg/cm² [b]	9.3	8.7	10.7	15.9	18.8	23.0	26.7	46.9

[a] General formula: $(13.6\pi r^2\alpha_z^2/2\pi r d 1000)p$.
[b] General formula: $(13.6\alpha_z\alpha_s r/1000d)p$.

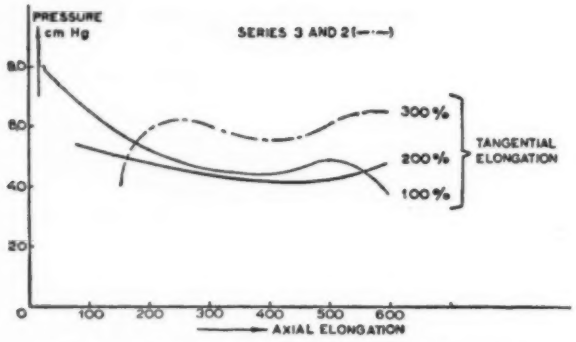


FIG. 3. Air pressure as a function of axial elongation.

It is clear from these figures that the tangential stress in all cases increases with axial elongation. The slope of this line increases with the transverse elongation.

Furthermore at a definite transverse elongation say 200 percent the tension must be equal to the stress parallel to the axis at an axial elongation of the same extent, in this case 200 percent. This is approximately true, however at 300 percent the referred stiffness of the tube caused somewhat higher values for the transverse stress.

RELAXATION

The decay with time of the stress parallel to the tube axis, was measured at the highest axial elongation i.e. 600 percent in the following way. The stress was measured at the moment of attaining the elongation of 600 percent and $\frac{1}{8}$—$\frac{1}{4}$—$\frac{1}{2}$—1—2—4—8—16 and 32 minutes after that.

A linear relationship is obtained if the stress is plotted *vs.* log time as is done for instance in Fig. 5.

The slope of this line $(F_{1/8}-F_{32}/-\log\frac{1}{8}+\log 32)$ is dependent on the original stress.

Instead of the original stress in our case the stress after one minute, F_1, is taken, as this value can be determined more accurately.

We calculated our relaxation constant from:

$$R_2=\frac{F_{1/4}-F_{16}}{6F_1}, \tag{2}$$

taking the number 2 as a base of our logarithmic system.

The relaxation constant R_2 was calculated for measurements made with tubes under variable conditions. The results are put together in Table II.

It is apparent from these data that the tangential elongation does not exert much influence on the relaxation constant though the actual tension is much increased. Nor is there any influence of the nature of the mix in so far that soft sulfur-vulcanized rubber and thiuram vulcanized rubber show the same relaxation constant. However, if a mix with a greater amount of sulfur (e.g. five parts) is used the relaxation constant falls to considerably lower values. This holds almost independently of the time of cure of this latter compound.

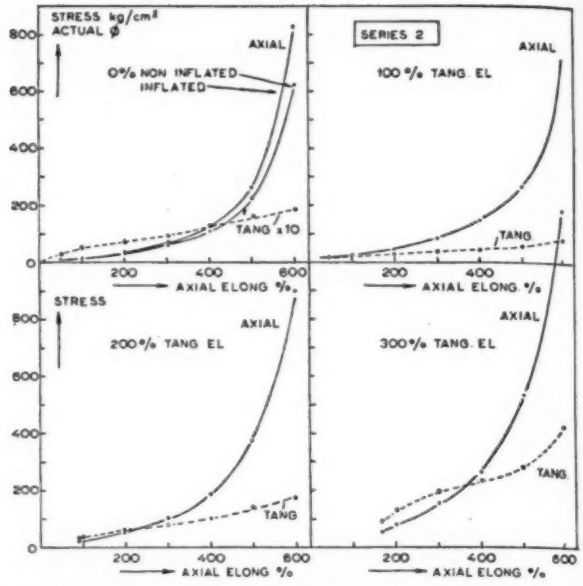


FIG. 4. Axial and tangential stress on actual cross section at biaxial elongation; rate of testing 30 cm per minute.

The relaxation constant is however sensitive to the axial elongation; at 500 percent the value is little more than half that at 600 percent elongation.

DISCUSSION

According to the now established statistical theory of elasticity[3] the elastic tension for unidirectional deformation is:

$$F=C[\alpha-(1/\alpha^2)]. \tag{3}$$

If one of the cross dimensions (e.g. the circumference of the tube) is kept constant this formula changes into

$$F_c=C[\alpha-(1/\alpha^3)]. \tag{4}$$

The following simple formulation of James and Guth's theory was kindly suggested by Ir. J. Bronkhorst of the Physical Department of the Rubber Foundation.

In Fig. 6, x, y, and z are dimensions of the rectangular sample. In

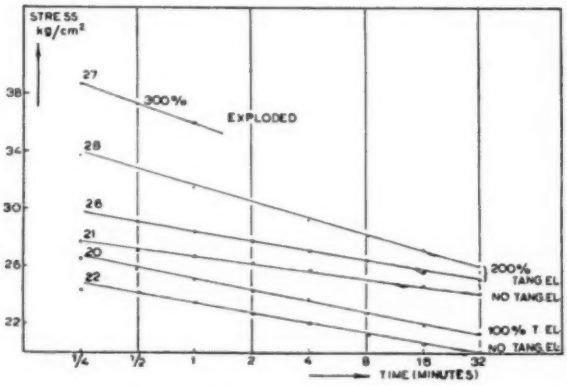


FIG. 5. Relaxation of stress with time.

[3] H. M. James and E. Guth, J. App. Phys. **15**, 294 (1944); J. Chem. Phys. **11**, 455 (1943).

unstrained condition $x=y=z=1$ cm. Then total molecular stress in x, y or z direction is $C=H$ (hydrostatic pressure in unstrained condition).

In a special strained condition x, y, z caused by forces F_x, F_y and F_z:

$F_x+Hyz=Cx=$ molecular tension
$F_y+Hxz=Cy$ $\quad x, y, z=1$ if the volume is considered constant.
$F_z+Hxy=Cz$

For unidirectional stress parallel to the x axis

$$F_y=F_z=0$$
$$Cy=Hxz$$
$$Cz=Hxy$$
$$F_x=Cx-Hyz=Cx-H/x$$
$$H=Cy/xz=C/x.$$
$$F_x=Cx-(C/x^2)$$

When z is kept constant by a force F_z, only $F_y=0$ (this is not exactly true as the inflation pressure represents a small negative value which is however neglected),

$$F_x+Hyz=Cx$$
$$Hxz=Cy$$
$$F_z+Hxy=Cz \quad H=Cy/xz=Cy^2=(C/x^2z^2).$$
$$F_x=Cx-Cy^2/x=C[x-(y^2/x)]=C[x-(1/z^2x^3)],$$
$$F_z=Cz-C(y^2/z)=C[z-(1/x^2z^3)].$$

Inflation to original diameter means $z=$ constant $=1$

$$F_x=C[x-(1/x^3)],$$
$$F_z=C[1-(1/x^2)].$$

Inflation to 100 percent transverse elongation ($z=2$):

$$F_x=C[x-(1/4x^3)],$$
$$F_z=C[2-(1/8x^2)].$$

Similarly at 200 percent transverse elongation:

$$\left.\begin{aligned}F_x&=C[x-(1/9x^3)]\\F_z&=C[3-(1/27x^2)]\end{aligned}\right\}.$$

At 300 percent transverse elongation:

$$\left.\begin{aligned}F_x&=C[x-(1/16x^3)]\\F_z&=C[4-(1/64x^2)]\end{aligned}\right\}.$$

For a definite elongation the second form is larger than the first one so it is to be expected that the stress-strain curve for a tube inflated to original cross section, will lie at higher stresses than the curve for the non-inflated tube. The difference should be relatively most pronounced at the lowest elongations.

The curves of Fig. 4 are not on a convenient scale to test these relations. In Table III several data are compared.

It appears from Table III that the axial stress of the inflated tube is relatively higher than would be expected from the theoretical formula. The stress of the non-inflated tubes at 100 percent and 200 percent corresponds with the less accurate values taken from dumbbells which amount to seven and ten kg/cm² respectively.

Similarly the ratio between axial stresses for non-inflated tubes and tubes inflated to 100 percent elongation is according to the theoretical derivation

$$\frac{\alpha-(1/\alpha^2)}{\alpha-(1/4\alpha^3)}\cdots. \qquad (5)$$

In Table IV the calculated and experimental values are compared.

TABLE II. Relaxation constants at 600 percent elongation.

	Rate of stretching 15 cm/min.	30 cm/min.
Series II		
Non-inflated	...	$R_2=0.017$
Inflated to original diameter	...	$R_2=0.021$
Inflated to 100 percent tangential elongation	...	$R_2=0.019$
Inflated to 200 percent tangential elongation	...	$R_2=0.023$
Inflated to 300 percent tangential elongation	...	$R_2=0.021$
Series III		
Non-inflated	...	$R_2=0.017$
Inflated to 100 percent tangential elongation	$R_2=0.023$	$R_2=0.019$
Inflated to 200 percent tangential elongation	...	$R_2=0.018$
Series V		
Non-inflated	0.022	$R_2=0.018$
Inflated to original diameter	0.021	$R_2=0.019$
Inflated to 100 percent tangential elongation	0.019	$R_2=0.020$
Inflated to 200 percent tangential elongation	0.017	$R_2=0.018$
Inflated to 300 percent tangential elongation	...	$R_2=0.019$
Dumbbells cut out of latex tubes	...	$R_2=0.017$
Dumbbells cut out of sheets vulcanized with 1.75 parts of sulfur, 20′ at 142°C in a press	...	$R_2=0.020$[b]
As above but vulcanized with three parts of tuads[a] instead of sulfur +accelerator	...	$R_2=0.020$[b]
As above, vulcanized with five parts of sulfur+accelerator	...	$R_2=0.014$
Mixture with carbon black	...	$R_2=0.021$

[a] Tetramethyl thiuramdisulphide.
[b] These vulcanizates show a value of $R_3=$ about 0.013 at 500 percent axis-elongation.

Though these figures are not very accurate they show a decided discrepancy with the calculated values; the experimental stresses being higher than those calculated.

This is even more obvious if we consider the transverse stress. Theory demands that the transverse stress, calculated on original cross section, shall remain practically constant at axial elongations above about 100 percent and under elongations where crystallinity disturbs the random configuration of the molecular network.

This is illustrated in Fig. 7 where the course of the

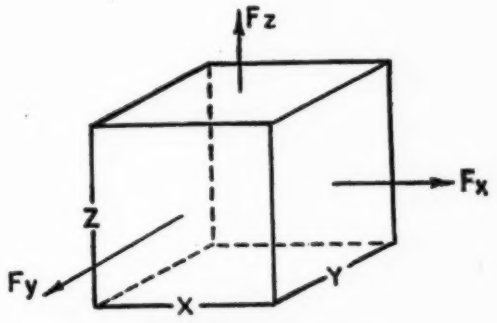


FIG. 6. Principle of stress.

TABLE III. Axial stresses on original cross section in kg/cm².

Series 2	Axial elongation 50 percent	100 percent	200 percent
Non-inflated	3.4	6.5	10.8
Inflated	4.4	7.9	12.3
Quotient	0.78	0.82	0.88
Quotient calculated from $\frac{[\alpha-(1/\alpha^2)]}{[\alpha-(1/\alpha^3)]}=$	0.87[5]	0.93[3]	0.97[5]

tangential stress is plotted schematically for a definite value of C (i.e., $C=3.3$). This value of C can be calculated from the stress-strain curve (non-inflated or dumbbell sample) as it is the axial stress where $\alpha-1/\alpha^2=1$ or $\alpha=$about 1.47 (47 percent elongation).

In our series 2 and 3 there is only approximate agreement at the very lowest elongations in axial or tangential direction. For instance at 0 percent tang. el., which means that the tube is inflated only to maintain its original diameter, the calculated curve only at the very beginning is approximately followed by the experiment. At 100 percent trans. el. however the first points up to about 100 percent axial elongation show some slight agreement, at higher elongations the difference increases strongly. It is to be expected that the constancy of volume assumed in theory will not hold exactly but it cannot explain the wide difference between experiment and theory.

Nor is it to be expected that neglecting of the small negative force in the third direction (the air pressure against the inner wall of the tube) would make much difference in the course of the theoretical curves.

So it seems that whereas there is reasonable agreement between experiment and theory for unidirectional stress this is no longer the case as soon as simultaneous elongation in a second direction is applied.

The high stresses above 300 percent axial elongation may point to the fact that no equilibrium is reached; if the tube decreases a little in diameter during axial stretching and afterwards is inflated to the correct diameter a higher pressure will be necessary than when both operations are accurately synchronized.

Treloar in his experiments about two-dimensional strain plotted the force per cm length of the extended sheet[4] (analogous to surface tension). As in his studies

TABLE IV. Calculated and experimental ratio of axial stress for non-inflated tubes and tubes inflated to 100 percent tangential elongation at various axial elongations.

Axial elongation	50 percent	100 percent	200 percent
Ratio calculated from.....(5)	0.74	0.89	0.97
Ratio measured in series 2	0.40	0.60	0.72
Ratio measured in series 3	0.52	0.71	0.80

[4] L. R. G. Treloar, Trans. Faraday Soc. **39**, 241 (1943).

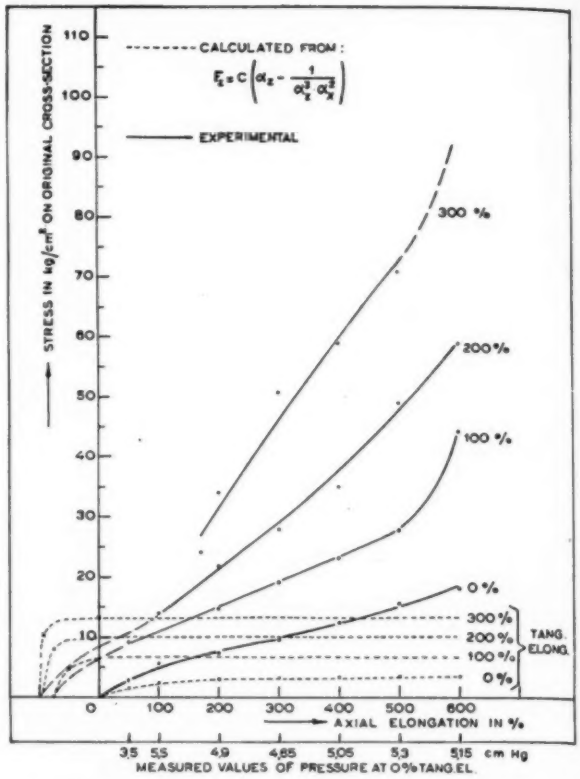


FIG. 7. Tangential stress as a function of axial elongation.

the strain in two dimensions was identical, the force per cm length of the sheet is

$$t=S_z/X^2=F_z/X=C\left(1-\frac{1}{X^6}\right)=C(1-\gamma^3), \quad (6)$$

when γ is the ratio of the final to the original thickness of the sheet and $X=Z$ the extension ratio in a direction parallel to the surface.

Treloar could show that the experimental tension force t at different extensions coincided with the theoretically derived equation

$$t=C(1-\gamma^3) \quad \text{for} \quad C=\text{about 4 kg/cm}^2$$

up to about 200 percent elongation ($\gamma=\frac{1}{3}$).

In our case the strain in the two principal directions is generally not the same but at every tangential elongation we have one moment where axial and tangential elongation are identical. In series V we have these cases at 56 percent, 100 percent, 200 percent, 275 percent and 300 percent elongation. At these points of identity of the two principal strains the stresses in the two perpendicular directions must be identical too; this appears to be the case with a deviation of about the experimental error of the stress measurements.

To compare our results with those of Treloar we have plotted the force t for these cases in Fig. 8 and it is evident from this figure that they agree fairly well with

theoretically calculated values for $C=5.7$ up to 200 percent elongation. This is shown in Table V.

So far there is agreement with theory as had been found by Treloar except that in our case the value of C is 5.7 against 4.0 in Treloar's experiments.

This however may be due to the different types of mixture used.**

In case the strains in the two perpendicular directions across the sheet are not identical we have two principal stresses and following again the method of plotting the entity t (stress per cm length of the sheet) we easily derive:

$$t_x=\frac{S_x}{XZ}=\frac{F_x}{Z}=C\left(\frac{X}{Z}-\frac{1}{X^3Z^3}\right)$$

and

$$t_z=\frac{S_z}{XZ}=\frac{F_z}{X}=C\left(\frac{Z}{X}-\frac{1}{X^3Z^3}\right)$$

for $X/Z=1$ this case simplifies again to the first problem. If we take however the strain situations, where X/Z is for instance $=2$ we would expect to find a similar curve as in Fig. 8 with the same value of C.

What actually is found Table VI illustrates. From Table VI it is evident that the C values do not agree with the C found for homogeneous two-dimensional strain and differ even when calculated from the axial and tangential stress respectively. So here also the deviations from theory are very pronounced. Mathematically these deviations may be accounted for partly by the treatment of Mooney[5] and more completely by that of Rivlin.[6]

The formula governing the relaxation which has been used here is

$$R=\frac{Ft_1-Ft_2}{Ft_1/(\log t_2-\log t_1)}; \qquad (7)$$

this expresses a constant decrease of stress in all time periods t_2-t_1, where $t_2=2t_1$ and t is counted from the moment the elongation was stopped. It is an approximation of the general formula of Tobolsky and Eyring.[7] It indicates that the decay of stress is the result of slippage of chain molecules along each other.

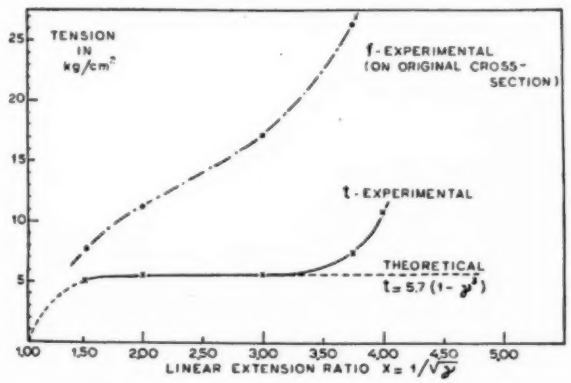


FIG. 8. Two-dimensional equilateral extension.

** For latex-rubber T. mentions a value of $C=6.3$.

[5] M. Mooney, J. App. Phys. **11**, 582 (1940).

[6] R. S. Rivlin, Phil. Trans. A, **240**, 459, 509 (1948); **241**, 379 (1948); **242**, 173 (1949); Proc. Roy. Soc. A, **195**, 463 (1949).

TABLE V. Forces at various equilateral elongations and the values of "C" calculated from these forces.

Elongation	F_x/X	"C"
56 percent	7.9/1.56=5.1	5.7
100 percent	11.2/2=5.6	5.6
200 percent	17.2/3=5.7	5.7
275 percent	26.5/3.75=7.1	7.1
300 percent	46/4=11.5	11.5

According to Flory[8] the relaxation results for the greater part from terminal chain slippage. This means that a rubber with longer primary molecules (before vulcanization) will have a lower relaxation constant than a rubber which has been thoroughly broken down to small molecular chains on the mixing mill.

If we compare two rubbers that have undergone the same milling treatment, so that we may assume about equal average molecular weight, but that have different degree of vulcanization, then the one with the highest cure will show the lower relaxation constant as the latter has the highest percentage of cross links and relatively the lowest number of terminal chains.

This is in agreement with the observation that the vulcanizate with five parts of sulfur shows a constant $R=0.014$ whereas the mixes vulcanized with 1.75 parts of sulfur and with thiuram only, show a value of $R=0.020$.

At these high elongations crystallization plays a part in stress-relaxation. The crystallites begin to form at about 250–300 percent elongation; this means that these first crystallites have a comparably low degree of

TABLE VI. Values of "C" calculated from axial and tangential stresses with $X/Z=2$.

Series V	Fx/Z	C calculated from Fx/Z	Fz/X	C calculated from Fz/X
Ax. el. 100 percent Tang. el. 0 percent	7.0– 7.5	3.7–4	2.6–3.2	6.9–8.6
Ax. el. 200 percent Tang. el. 50 percent	9.1	4.6	3.6	7.2
Ax. el. 300 percent Tang. el. 100 percent	10.1–11.8	5.1–5.9	3.9–4.0	8.0
Ax. el. 400 percent Tang. el. 150 percent	14.2–15.8	7.2–7.9		
Ax. el. 500 percent Tang. el. 200 percent	17.2–18.5	8.6–9.3	4.0–5.0	8–10

[7] A. V. Tobolsky and H. Eyring, J. Chem. Phys. **11**, 125 (1943). The simplified formula was experimentally determined by P. Phillips, Proc. Phys. Soc. (London) **19**, 491 (1905).

[8] P. J. Flory, Ind. Eng. Chem. **38**, 417 (1946).

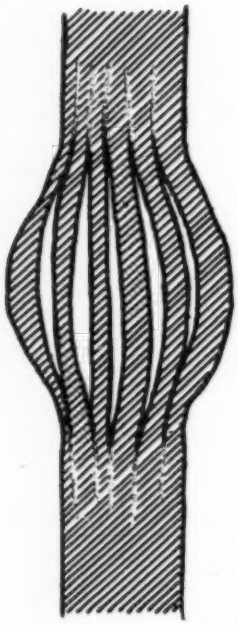


FIG. 9. Effect of inflation on a stretched sample.

orientation, lower than those formed at 400–500 percent elongation and these again are less orientated than crystals formed at 500–600 percent elongation.

In the latter state there are many crystallites that were in equilibrium with the surrounding amorphous mass at elongations much lower than 600 percent.

These strive to reach an equilibrium at the now (600 percent el.) prevailing conditions by re-crystallization and detaching of terminal chains from crystallites. It is probable that such a re-crystallization is a rather slow process as compared with the slippage of terminal chains but it is difficult to distinguish between them.

Mooney *et al.*[9] besides, report indications of a relaxation mechanism of a higher rate than that found in normal analysis; its existence was established on Shore Durometer readings $\frac{1}{10}$ sec. after stretching.

In our case however we would expect that transverse strain would severely hamper orientation and crystallization in axis parallel direction and diminish its contribution to stress-relaxation. The effect on crystallization can be experimentally shown by first strongly elongating one of our tube-samples and then inflating it. Very high air pressure is necessary to cause the first transverse elongation which then does not turn out homogeneously but occurs by a definite tearing of the opaque crystallized material into streaks of thinner and clear amorphous areas and remaining crystallized fiber-like bundles. This situation is schematically drawn in Fig. 9.

There has not been found definite influence of the transverse strain on the axial relaxation. It may be that the measurements, which are technically difficult to perform and absorb much time, have not been sufficiently accurate to distinguish the differences but then the effects could not be very pronounced.

[9] Mooney, Wolstenhome, and Villars, J. App. Phys. **15**, 324 (1944).

CONCLUSIONS

(1) The stress-strain curves of vulcanized rubber held at definite tangential elongations during stretching differ more from the simple unidirectional-extension-stress curve than the theory predicts. The differences are larger at higher tangential elongations.

(2) The relaxation constant determined at 600 percent elongation is practically unaffected by tangential elongations up to 300 percent. The investigations are part of the work on fundamental properties of rubber of the Research Department of the Rubber-Stichting at Delft.

On the Endurance Limit of a Round Bar with Longitudinal Grooves

H. Ōkubo
Institute of High Speed Mechanics, Tōhoku University, Sendai, Japan
(Received June 2, 1950)

In a previous paper, we reached a consideration that the endurance limit of a material should be determined by the amount of the mean stress in a certain area around a point, instead of the stress at the point, and the extent of the area is determined for several kind of metals, by using the results of tension-compression fatigue tests.

In this paper, the extent of area where the mean value of stress to be taken, is determined for a mild steel in another way, viz., by a torsion fatigue test, and we reached the conclusion that the extent of area chiefly depends on the kind of metal and independent of the course of the experiment.

IN our previous study, the influence of the diameterical hole on the endurance limit of round bars has been investigated.[1] The theory of elasticity shows that the stress concentration around a diameterical hole of a round bar under tension, increases as the diameter of hole decreases and when the diameter of hole decreases infinitely, the stress concentration factor approaches 3.[2] While, according to our experimental results for several kinds of metals, with decrease in the diameter of the hole, the endurance limit increases, when the diameter of the hole is very small. The theory fails to explain the experimental results. The discrepancy between the theoretical and the experimental results arises from a fact that the theory assumes the material to be perfectly homogeneous and isotropic, while the metals in practical use are neither homogeneous nor isotropic. To correct the theory so as to meet the experimental results, we reached the consideration that when we determine the endurance limit, we must consider the mean stress in a certain area around the point, instead of the stress at a point. To simplify the calculation, we took up the mean value of the stress in a certain line, in place of the mean value in the area. The magnitude of the distance h where the mean value is taken, was considered to depend chiefly on the size of crystals of the metal and so it is a constant, determination of which is pending on the kind of metals. Under the above considerations, the values of h for some metals, determined by using our previous fatigue data, were

Mild steel	0.23 mm
Special steel	0.15 mm
Super-duralumin	0.13 mm

Our experiments for these metals were done by means of tension-compression tests using Haigh's testing machine.

In this paper, we shall determine the constant h for a mild steel in another way, *viz.*, by means of a torsion fatigue test, using specimens of round bars with longitudinal grooves as shown in Fig. 1. The specimens are cut out very carefully and the dimensions of grooves checked by a plate gauge which is measured accurately by a Zeiss comparator.

The mechanical properties of the steel are given in Table I.

The testing machine, used in our present experiment is of mechanical type. The specimen is placed in series with a weigh-bar, whose stress-strain relation has been previously obtained by statical loading experiments. The torque acting on the specimen is estimated by observing the twist of the weigh-bar and its twist is measured optically by means of a mirror placed at its moving end.

The scale is so graduated as to allow the direct reading of the maximum shearing stress for specimens of the diameter 10 mm. The fatigue data obtained in connection with the above mentioned program are given in Tables II and III. Figure 2 shows a fractured specimen with longitudinal grooves. The endurance limit of the steel is 11.5 Kg/mm².[3]

The fatigue cracks, appeared on the bottoms of grooves of specimens for different loadings, are shown in Fig. 3. When the magnitude of the reversed stress slightly passes over the endurance limit, very small cracks, as shown in Fig. 3(a), begin to appear locally on the bottom of the groove. According as the stress increases, the small cracks appear at every part of the bottom and join together, and then some of the enlarged cracks spread along the bottom of the groove (Fig. 3(b)), and finally the specimen is broken. We shall define the endurance limit for the specimens with grooves, by the

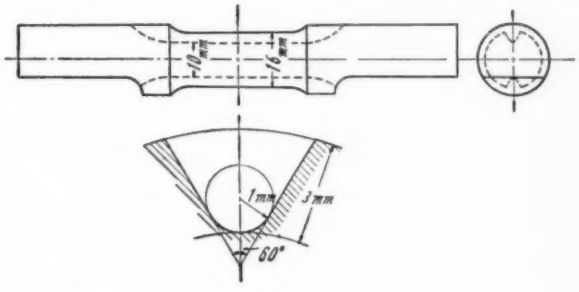


FIG. 1. Test piece.

[1] K. Minamiōzi and H. Ōkubo, J. Franklin Inst. **249**, 49 (1950).
[2] R. E. Peterson, J. App. Mech. (ASME), p. *A*-15 (1936).
[3] For the endurance limit of specimens without grooves, the maximum stress is taken, since the variation in the interval h is very small.

FIG. 2.

limiting value of the stress, below which cracks will not occur. Then, the reading corresponding to the endurance limit for the specimens with grooves will be determined from Table III, as 17.3 Kg/mm².

Next, we shall make some calculations to find the stress distribution in the specimen corresponding to the reading.

Let us consider a circular cylinder of the cross section as shown in Fig. 4, and take the axis of z to be parallel to the generators of the cylinder. Then the stresses which do not vanish are

$$X_z=\mu\tau\left(\frac{\partial\phi}{\partial y}-y\right),\quad Y_z=-\mu\tau\left(\frac{\partial\phi}{\partial x}-x\right),\qquad(1)$$

where μ and τ represent the modulus of rigidity and the twist respectively. The function ϕ satisfies the equation

$$\partial^2\phi/\partial x^2+\partial^2\phi/\partial y^2=0,\qquad(2)$$

at all points within the bounding curve of the cross section, and at the boundary the condition

$$\phi-\tfrac{1}{2}(x^2+y^2)=\text{const.}\qquad(3)$$

We shall consider a function ϕ in the form[4]

$$\phi=\frac{(1-r^4)A_0}{1+r^4-2r^2\cos2\theta}+A_1\left\{\frac{1-a_1^2r^2\cos2\theta}{1+a_1^4r^4-2a_1^2r^2\cos2\theta}-\frac{r^4-a_1^2r^2\cos2\theta}{a_1^4+r^4-2a_1^2r^2\cos2\theta}\right\}.\qquad(4)$$

TABLE I.

Yield strength Kg/mm²	Ultimate strength Kg/mm²	True breaking strength Kg/mm²	Elongation percent	Reduction of area percent
25.5 (upper) 22.3 (lower)	38.2	83.0	36.6	67.9

TABLE II. Specimens of 10-mm diameter without grooves.

Specimen No.	Range of stress Kg/mm²	No. of repetitions	Remarks
1	±13.0	0.106×10^7	broken
2	±12.6	0.392×10^7	broken
3	±12.0	0.596×10^7	broken
4	*±11.5*	1.003×10^7	unbroken

[4] H. Ōkubo, "On the torsion of a shaft with keyways," Quart. J. Mech. App. Math. (to be published).

TABLE III. Specimens of 16-mm diameter with longitudinal grooves, as shown in Fig. 1.

Specimen No.	Reading of the nominal range of stress Kg/mm²	No. of repetitions	Remarks
1	±43.2	0.037×10^7	broken
2	±36.3	0.292×10^7	broken
3	±33.1	0.724×10^7	broken
4	*±32.7*	1.097×10^7	unbroken, cracked
5	±25.5	1.043×10^7	unbroken, cracked
6	±22.8	1.000×10^7	unbroken, cracked
7	±22.1	1.320×10^7	unbroken, cracked
8	±19.2	1.002×10^7	unbroken, cracked
9	±18.3	1.100×10^7	unbroken, cracked
10	±17.8	1.019×10^7	unbroken, cracked
11	*±17.3*	1.000×10^7	uncracked

The above function satisfies the Eq. (2), and is singular at the two points ($\theta=0,\ \pi$) on the bounding circle and at the two points ($r=a_1,\ \theta=0,\ \pi$) within the circle (a_1 being assumed to be smaller than unity).

From Eqs. (3) and (4), the equation which represents the boundary of the cross section is

$$\frac{(1-r^4)A_0}{1+r^4-2r^2\cos2\theta}+A_1\left\{\frac{1-a_1^2r^2\cos2\theta}{1+a_1^4r^4-2a_1^2r^2\cos2\theta}-\frac{r^4-a_1^2r^2\cos2\theta}{a_1^4+r^4-2a_1^2r^2\cos2\theta}\right\}-\tfrac{1}{2}r^2=-\tfrac{1}{2}.\qquad(5)$$

Equation (5) represents a circle of radius unity and the two curves represented by the equation:

$$\frac{A_0}{1+r^4-2r^2\cos2\theta}+\frac{A_1\{a_1^4(1+r^4)-a_1^2(1+a_1^4)r^2\cos2\theta\}}{(1+a_1^4r^4-2a_1^2r^2\cos2\theta)(a_1^4+r^4-2a_1^2r^2\cos2\theta)}+\frac{1}{2(1+r^2)}=0.\qquad(6)$$

When the curve represented by Eq. (6), agrees almost with the shape of the given notch, then we can use the Eq. (4), as an approximate solution for the problem.

To determine the unknown constants A_0 and A_1 in Eq. (4), we shall assume that the curve represented by Eq. (6) passes through the two points ($r=0.625,\ \theta=0°$ and $r=0.81623,\ \theta=12°10'$) on the periphery of the notch, then assuming $a_1=0.7562$, we have

$$A_0=-0.065507,\quad A_1=-0.045069.$$

If we denote the radius of curvature at the bottom of the notch by ρ, then

$$\rho=\frac{b^2}{b-(d^2r/d\theta^2)_0},\qquad(7)$$

where

$$\left(\frac{d^2r}{d\theta^2}\right)_0=-\frac{[8A_0b^2(1+b^2)/(1-b^2)^3]+4A_1a_1^2b^2\{[1+a_1^2b^2/(1-a_1^2b^2)^3]+[a_1^2+b^2/(a_1^2-b^2)^3]\}}{b-[4A_0b/(1-b^2)^2]-2A_1a_1^2b\{[1/(1-a_1^2b^2)^2]+[1/(a_1^2-b^2)^2]\}},$$

and b is the distance from the center of the cross section to the bottom of a notch.

Putting $b=0.625$, it becomes $\rho=0.1250$. The curve represented by Eq. (6), is shown by a dotted line in Fig. 4, which agrees practically with the given shape of the notch as is shown in the figure. Accordingly, using the solution given by Eq. (4), we shall proceed with our further calculations.

The stress on the x axis is calculated from Eq. (1), as

$$Y_z=\mu\tau\left\{r-\frac{4A_0r}{(1-r^2)^2}-2A_1a_1^2r\left(\frac{1}{(1-a_1^2r^2)^2}+\frac{1}{(a_1^2-r^2)^2}\right)\right\}, \quad (8)$$

and so the mean value of the stress from the bottom of the notch to a distance h along the x axis is

$$\bar{Y}_z=\frac{1}{h}\int_{b-h}^{b}Y_z\,dr=\frac{\mu\tau}{h}\left|\frac{r^2}{2}-\frac{2A_0}{1-r^2}-A_1\left(\frac{1}{1-a_1^2r^2}+\frac{a_1^2}{a_1^2-r^2}\right)\right|_{b-h}^{b}. \quad (9)$$

The value of $\bar{Y}_z$ varies with the magnitude of the distance h, and the relation between $\bar{Y}_z$ and h, calculated by Eq. (9), is shown in Fig. 5. The torque acting on the cylinder is[5]

$$T=\mu\tau\left(I-\oint\psi\frac{\partial\phi}{\partial s}ds\right), \quad (10)$$

where I is the polar moment of inertia of the cross section about z axis, ψ is a function conjugate with ϕ and given by

$$\psi=-\frac{2A_0r^2\sin2\theta}{1+r^4-2r^2\cos2\theta}-A_1a_1^2r^2\sin2\theta\left\{\frac{1}{1+a_1^4r^4-2a_1^2r^2\cos2\theta}+\frac{1}{a_1^4+r^4-2a_1^2r^2\cos2\theta}\right\},$$

and the integral is taken around the bounding curve of the cross section.

For the section, shown in Fig. 4,

$$I=1.398 \quad \text{and} \quad \oint\psi\frac{\partial\phi}{\partial s}ds=0.311,$$

and hence

$$T=1.087\mu\tau.$$

The torque T_0 acting on a cylinder of radius 0.625 without grooves, when twisted by the amount τ_0 is

$$T_0=0.2397\mu\tau_0.$$

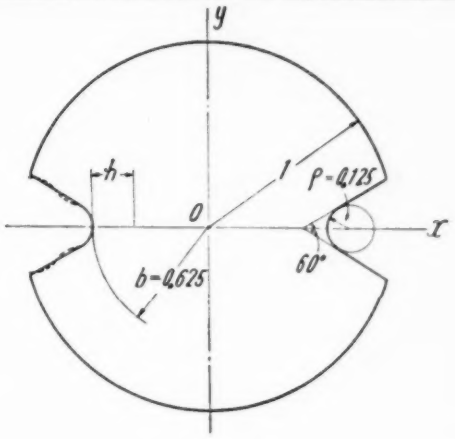
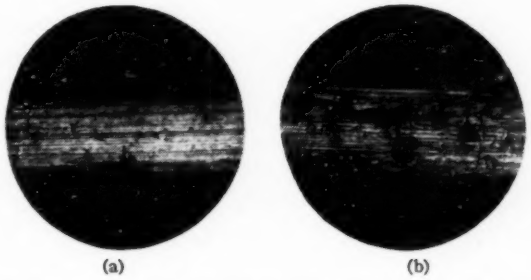


FIG. 4.



(a) (b)

FIG. 3. Micro-photographs of the cracks (×100). (a) Reading of the nominal range of stress, ±17.8 Kg/mm². No. of repetitions =1.019×10⁷. (b) Reading of the nominal range of stress, 25.5 Kg/mm². No. of repetitions=1.043×10⁷.

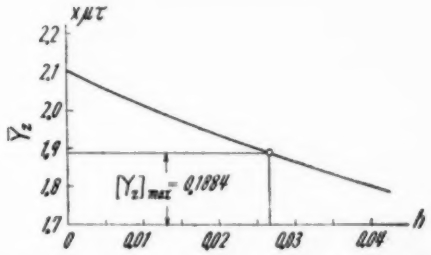


FIG. 5. The relation between the mean stress and the interval h.

[5] H. Love, *The Mathematical Theory of Elasticity* (1927) 4th Ed., p. 312.

Hence, we have

$$T_0/T=0.2205(\tau_0/\tau).$$

According to our fatigue data for a cylinder of radius 5 mm, without grooves, the endurance limit of the metal is 11.5 Kg/mm². While the reading corresponding to the endurance limit for a cylinder of radius 8 mm with grooves is 17.3 Kg/mm². The reading does not denote directly the maximum stress but represents the value proportional to the torque acting on the specimen, when the dimensions of the specimen differ from the standard type (a cylinder of radius 5 mm without grooves), subsequently

$$T_0/T=11.5/17.3,$$

and hence

$$0.2205(\tau_0/\tau)=11.5/17.3,$$

or

$$\tau_0=3.015\tau.$$

The maximum stress arises in a cylinder of radius 0.625, when twisted by the amount τ_0 is

$$[Y_z]_{\max}=0.625\mu\tau_0=1.884\mu\tau.$$

The endurance limit of a metal, however, is a constant, which is independent of the form or the size of the specimen. Subsequently, $[Y_z]_{\max}$ must be equal to $\bar{Y}_z$, by virtue of our previous assumption that for the determination of the endurance limit of specimens, the mean value of stress in the interval h must be taken up instead of the maximum stress.

If we determine h from our fatigue data, so as to satisfy the above mentioned condition (see Fig. 5), it becomes

$$h=0.0265,$$

or

$$h=0.21 \text{ mm}.$$

The value of h for a mild steel, previously obtained by using our tension-compression fatigue data is

$$h=0.23 \text{ mm},$$

and the results almost agree in two cases.

From this result, we can conclude that the value of h depends only on the kind of metal, and independent of the course of the experiment.

ACKNOWLEDGMENT

A grant in aid for scientific research of the Education Ministry of Japan has been given to the present study.

The author also takes this opportunity to thank his assistants, T. Uematsu and E. Itagaki, for their help in this study.

The Magnetic Field of a Plane Circular Loop

C. L. BARTBERGER

Naval Air Experimental Station, Philadelphia, Pennsylvania

(Received March 9, 1950)

The axial and radial components of the magnetic field of a plane circular loop are expressed in terms of cylindrical coordinates. The expressions involve two integrals which are related to certain of the complete elliptic integrals. Tables of values of these integrals are presented. Interpolation in these tables facilitates rapid calculation of the field components.

INTRODUCTION

THE problem of calculating the radial and axial components of the magnetic field produced by the current in a plane circular loop of wire is well known and has been treated in numerous textbooks.[1,2] Recently Blewett,[3] making use of the solution in terms of cylindrical coordinates and complete elliptic integrals, has computed and tabulated the values of the field components of numerous points in the vicinity of the loop. Such a tabulation, however, has the difficulty of being two-dimensional, since the field is a function of both the axial and radial coordinates. It is therefore not practical to tabulate the field at a sufficient number of points to permit easy interpolation.

A few years ago the need arose to determine the magnetic field produced by a pair of Helmholtz coils which were used for testing aircraft compasses. The solution was carried out in cylindrical coordinates by a slightly different method from that used by Blewett. The resulting expressions for the field components are relatively simple and involve two integrals which are related to certain of the complete elliptic integrals. Once these integrals have been evaluated, the computation of the field components to high accuracy becomes a simple matter. It is the purpose of this paper

[1] J. H. Jeans, *The Mathematical Theory of Electricity and Magnetism*, 5th Edition (Cambridge University Press, London, 1925), pp. 431, 432.

[2] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1939), pp. 267 and 270.

[3] J. P. Blewett, "Magnetic field configurations due to air core coils," J. App. Phys. 18, 968 (1947).

to publish tables of values of the two integrals together with formulas for their use. Values are given for sufficiently small increments of the argument to permit linear interpolation throughout the greater part of the tables.

MAGNETIC FIELD COMPONENTS DUE TO CURRENT IN A PLANE CIRCULAR LOOP

If z and ρ represent axial and radial coordinates measured from the center of the loop, the axial and radial components of the magnetic field can be expressed as follows:

$$H_z=(H_0/s^3)[I_1-(\rho/a)I_2] \tag{1}$$

$$H_\rho=(H_0 z/s^3 a)I_2, \tag{2}$$

where a = radius of the loop, $s^2=(a^2+z^2+\rho^2)/a^2$, and H_0 = field at center of loop. The integrals I_1 and I_2 are

$$I_1=\frac{1}{\pi}\int_0^\pi \frac{d\theta}{(1-b\cos\theta)^{\frac{3}{2}}} \tag{3}$$

$$I_2=\frac{1}{\pi}\int_0^\pi \frac{\cos\theta d\theta}{(1-b\cos\theta)^{\frac{3}{2}}}, \tag{4}$$

where

$$b=\frac{2a\rho}{a^2+z^2+\rho^2}. \tag{5}$$

The integrals I_1 and I_2 can be expressed in terms of complete elliptic integrals if the following substitutions are made:

$$\phi=\tfrac{1}{2}(\pi-\theta) \tag{6}$$

$$k^2=2b/(1+b). \tag{7}$$

The results are

$$I_1=\frac{2}{\pi}\frac{(1-k^2/2)^{\frac{1}{2}}}{1-k^2}E \tag{8}$$

$$I_2=\frac{2}{\pi}\frac{(1-k^2/2)^{\frac{1}{2}}}{1-k^2}(2B-E), \tag{9}$$

where

$$E=\int_0^{\pi/2}(1-k^2\sin^2\phi)^{\frac{1}{2}}d\phi \tag{10}$$

and

$$B=\int_0^{\pi/2}\frac{\cos^2\phi d\phi}{(1-k^2\sin^2\phi)^{\frac{1}{2}}}. \tag{11}$$

SERIES EXPRESSIONS FOR INTEGRALS

If the integrands are expanded in series and integrated term by term, the following expressions are obtained for I_1 and I_2:

$$I_1=1+\frac{3\cdot5}{4\cdot4}b^2+\frac{3\cdot5\cdot7\cdot9}{4\cdot4\cdot8\cdot8}b^4+\frac{3\cdot5\cdot7\cdot9\cdot11\cdot13}{4\cdot4\cdot8\cdot8\cdot12\cdot12}b^6+\cdots \tag{12}$$

$$I_2=\frac{3}{4}b+\frac{3\cdot5\cdot7}{4\cdot4\cdot8}b^3+\frac{3\cdot5\cdot7\cdot9\cdot11}{4\cdot4\cdot8\cdot8\cdot12}b^5+\cdots. \tag{13}$$

For values of b near unity these series converge slowly. With the aid of expressions for E and B in the neighborhood of $k=1$, I_1 and I_2 can be expressed as series in powers of $(1-b)$. Let

$$k'^2=1-k^2$$

Then[4]

$$E=1+\sum_{n=1}^{\infty}\frac{1^2\cdot3^2\cdots(2n-3)^2(2n-1)}{2^2\cdot4^2\cdots(2n-2)^2 2n}(\Lambda-e_n)k'^{2n} \tag{14}$$

$$B=1-\sum_{n=1}^{\infty}\frac{1^2\cdot3^2\cdots(2n-3)^2(2n-1)^2}{2^2\cdot4^2\cdots(2n-2)^2 2n}(\Lambda-b_n)k'^{2n}, \tag{15}$$

where

$$e_n=\frac{1}{2n}\left[1+\frac{6n-2}{1\cdot1\cdot3}+\frac{10n-4}{2\cdot3\cdot5}+\cdots+\frac{(2n-1)2n-(2n-2)}{(n-1)(2n-3)(2n-1)}\right] \tag{16}$$

$$b_n=\frac{1}{2n}\left[2n+1+\frac{2n-2}{1\cdot1\cdot3}+\frac{2n-4}{2\cdot3\cdot5}+\cdots+\frac{2}{(n-1)(2n-3)(2n-1)}\right] \tag{17}$$

$\Lambda=\ln 4/k'$.

The integrals I_1 and I_2 become

$$I_1=\frac{\sqrt{2}}{\pi}\left\{\left[\frac{1}{1-b}+\frac{1}{8}-\frac{13}{256}(1-b)-\frac{103}{1024}(1-b)^2-\frac{86567}{786432}(1-b)^3-\cdots\right]+\Lambda\left[\frac{1}{4}+\frac{15}{64}(1-b)+\frac{195}{1024}(1-b^2)+\frac{4935}{32768}(1-b)^3+\cdots\right]\right\} \tag{18}$$

$$I_2=\frac{\sqrt{2}}{\pi}\left\{\left[\frac{1}{1-b}+\frac{9}{8}+\frac{307}{256}(1-b)+\frac{1263}{1024}(1-b)^2+\frac{327219}{262144}(1-b)^3+\cdots\right]\right.$$

[4] F. Jahnke and E. Emde, *Tables of Functions* (Dover Publications, New York, 1933), pp. 73.

$$-\Lambda\left[\frac{3}{4}+\frac{57}{64}(1-b)+\frac{945}{1024}(1-b)^2+\frac{30345}{32768}(1-b)^3+\cdots\right]\Big\} \quad (19)$$

$\Lambda=\ln 4-\frac{1}{2}\ln[(1-b)/(1+b)]$.

THE TABLES

Table I lists values of I_1 for values of b in increments of 0.001 from 0 to 0.800. Table II lists corresponding values of I_2. Table III lists values of I_1, I_2, I_1-I_2, $(1-b)I_1$, and $(1-b)I_2$ for values of b in increments of 0.001 from 0.800 to 1.000. Table IV lists values of the same five functions for values of b in increments of 0.0001 from 0.9950 to 1.0000. The difference I_1-I_2 is included to facilitate the computation of H_z in the vicinity of $b=1$ (near the wire). In this case Eq. (1) involves the subtraction of two large quantities which are nearly equal. This difficulty can be circumvented by using I_1-I_2 as follows:

$$H_z=(H_0/s^3)\{(I_1-I_2)+[1-(\rho/a)]I_2\}.$$

The functions $(1-b)I_1$ and $(1-b)I_2$ are of value in computing I_1 and I_2 in the vicinity of $b=1$, since they lend themselves to linear interpolation.

TABLE I. Values of the integral I_1.

b	0	1	2	3	4	5	6	7	8	9
0.00	1.000 000	000 001	000 004	000 008	000 015	000 023	000 034	000 046	000 060	000 076
0.01	000 094	000 114	000 135	000 159	000 184	000 211	000 240	000 271	000 304	000 339
0.02	000 375	000 414	000 454	000 496	000 540	000 586	000 634	000 684	000 736	000 789
0.03	000 844	000 902	000 961	001 022	001 085	001 150	001 217	001 285	001 356	001 428
0.04	001 502	001 579	001 657	001 737	001 818	001 902	001 988	002 075	002 165	002 256
0.05	1.002 349	002 445	002 542	002 641	002 742	002 844	002 949	003 056	003 164	003 275
0.06	003 387	003 501	003 618	003 736	003 856	003 978	004 101	004 227	004 355	004 484
0.07	004 616	004 749	004 885	005 022	005 162	005 303	005 446	005 591	005 738	005 887
0.08	006 038	006 191	006 346	006 503	006 661	006 822	006 985	007 149	007 316	007 484
0.09	007 655	007 827	008 002	008 178	008 356	008 537	008 719	008 903	009 090	009 278
0.10	1.009 468	009 660	009 855	010 051	010 249	010 449	010 652	010 856	011 062	011 270
0.11	011 480	011 693	011 907	012 123	012 342	012 562	012 784	013 009	013 235	013 464
0.12	013 694	013 927	014 161	014 398	014 636	014 877	015 120	015 365	015 612	015 861
0.13	016 112	016 365	016 620	016 877	017 137	017 398	017 662	017 927	018 195	018 465
0.14	018 737	019 011	019 287	019 565	019 845	020 128	020 412	020 699	020 988	021 278
0.15	1.021 572	021 867	022 164	022 464	022 765	023 069	023 375	023 683	023 994	024 306
0.16	024 621	024 937	025 256	025 577	025 901	026 226	026 554	026 884	027 216	027 551
0.17	027 887	028 226	028 567	028 910	029 256	029 604	029 954	030 306	030 660	031 017
0.18	031 376	031 737	032 101	032 466	032 835	033 205	033 578	033 953	034 330	034 709
0.19	035 091	035 475	035 862	036 251	036 642	037 035	037 431	037 829	038 230	038 632
0.20	1.039 038	039 445	039 855	040 267	040 682	041 099	041 519	041 941	042 365	042 792
0.21	043 221	043 652	044 086	044 523	044 961	045 403	045 846	046 293	046 741	047 192
0.22	047 646	048 102	048 561	049 022	049 485	049 951	050 420	050 891	051 365	051 841
0.23	052 320	052 801	053 285	053 771	054 260	054 751	055 245	055 742	056 241	056 743
0.24	057 248	057 755	058 264	058 777	059 292	059 809	060 330	060 852	061 378	061 906
0.25	1.062 437	062 971	063 507	064 046	064 588	065 132	065 680	066 230	066 782	067 338
0.26	067 896	068 457	069 020	069 587	070 156	070 728	071 303	071 881	072 462	073 045
0.27	073 631	074 220	074 812	075 407	076 005	076 605	077 209	077 815	078 424	079 036
0.28	079 651	080 269	080 890	081 514	082 141	082 771	083 404	084 040	084 679	085 321
0.29	085 966	086 614	087 265	087 919	088 576	089 236	089 900	090 566	091 236	091 908
0.30	1.092 584	093 263	093 945	094 630	095 319	096 010	096 705	097 403	098 104	098 808
0.31	099 516	100 227	100 941	101 658	102 379	103 103	103 830	104 561	105 295	106 032
0.32	106 773	107 517	108 264	109 015	109 769	110 526	111 287	112 052	112 820	113 591
0.33	114 366	115 144	115 926	116 711	117 500	118 292	119 088	119 887	120 690	121 497
0.34	122 307	123 121	123 939	124 760	125 585	126 413	127 245	128 081	128 921	129 764
0.35	1.130 611	131 462	132 316	133 175	134 037	134 903	135 773	136 647	137 524	138 406
0.36	139 291	140 180	141 073	141 971	142 872	143 777	144 686	145 599	146 516	147 437
0.37	148 362	149 291	150 225	151 162	152 104	153 050	153 999	154 953	155 912	156 874
0.38	157 841	158 812	159 787	160 766	161 750	162 738	163 731	164 727	165 728	166 734
0.39	167 744	168 758	169 777	170 800	171 828	172 861	173 898	174 939	175 985	177 035
0.40	1.178 091	179 150	180 215	181 284	182 358	183 436	184 520	185 608	186 700	187 798
0.41	188 900	190 008	191 120	192 237	193 359	194 485	195 617	196 754	197 896	199 042
0.42	200 194	201 351	202 513	203 680	204 853	206 030	207 213	208 401	209 594	210 792
0.43	211 996	213 205	214 419	215 639	216 864	218 094	219 330	220 572	221 818	223 071
0.44	224 329	225 592	226 862	228 136	229 417	230 703	231 995	233 292	234 596	235 905

TABLE I.—*Continued.*

b	0	1	2	3	4	5	6	7	8	9
0.45	1.237 220	238 541	239 868	241 200	242 539	243 884	245 234	246 591	247 954	249 323
0.46	250 698	252 079	253 466	254 860	256 260	257 666	259 079	260 497	261 923	263 354
0.47	264 793	266 237	267 689	269 146	270 611	272 082	273 559	275 044	276 535	278 033
0.48	279 538	281 049	282 568	284 093	285 626	287 165	288 711	290 265	291 826	293 393
0.49	294 968	296 551	298 140	299 737	301 341	302 953	304 572	306 198	307 833	309 474
0.50	1.311 124	312 780	314 445	316 117	317 798	319 486	321 181	322 885	324 597	326 317
0.51	328 045	329 781	331 525	333 277	335 038	336 807	338 584	340 369	342 163	343 966
0.52	345 777	347 597	349 425	351 262	353 108	354 963	356 826	358 699	360 580	362 471
0.53	364 370	366 279	368 197	370 124	372 060	374 006	375 961	377 926	379 900	381 884
0.54	383 877	385 880	387 893	389 916	391 949	393 991	396 044	398 107	400 180	402 263
0.55	1.404 356	406 460	408 574	410 699	412 834	414 980	417 137	419 304	421 482	423 671
0.56	425 871	428 083	430 305	432 538	434 783	437 039	439 306	441 585	443 876	446 178
0.57	448 492	450 818	453 156	455 505	457 867	460 241	462 627	465 025	467 436	469 859
0.58	472 294	474 743	477 204	479 678	482 165	484 665	487 177	489 704	492 243	494 796
0.59	497 362	499 942	502 535	505 143	507 764	510 399	513 048	515 711	518 389	521 081
0.60	1.523 787	526 508	529 244	531 995	534 760	537 540	540 336	543 147	545 973	548 814
0.61	551 672	554 544	557 433	560 338	563 258	566 195	569 149	572 118	575 104	578 107
0.62	581 127	584 164	587 218	590 289	593 377	596 483	599 606	602 747	605 906	609 084
0.63	612 279	615 493	618 725	621 975	625 245	628 533	631 841	635 167	638 513	641 879
0.64	645 264	648 670	652 095	655 540	659 006	662 493	666 000	669 528	673 077	676 647
0.65	1.680 238	683 852	687 487	691 143	694 822	698 524	702 248	705 994	709 764	713 557
0.66	717 372	721 212	725 075	728 962	732 874	736 809	740 770	744 754	748 764	752 800
0.67	756 860	760 947	765 059	769 197	773 362	777 554	781 772	786 017	790 290	794 591
0.68	798 919	803 275	807 660	812 074	816 516	820 988	825 489	830 020	834 580	839 172
0.69	843 793	848 446	853 130	857 845	862 593	867 372	872 184	877 028	881 906	886 817
0.70	1.891 762	896 740	901 754	906 802	911 885	917 003	922 158	927 348	932 575	937 839
0.71	943 140	948 479	953 856	959 271	964 725	970 219	975 752	981 325	986 939	992 594
0.72	998 290	*004 027	*009 808	*015 630	*021 496	*027 406	*033 360	*039 358	*045 401	*051 490
0.73	2.057 625	063 807	070 036	076 312	082 636	089 009	095 432	101 904	108 427	115 000
0.74	121 625	128 302	135 032	141 815	148 652	155 544	162 491	169 494	176 553	183 670
0.75	2.190 844	198 077	205 370	212 722	220 135	227 609	235 146	242 745	250 409	258 136
0.76	265 930	273 789	281 715	289 709	297 772	305 904	314 107	322 381	330 728	339 147
0.77	347 641	356 209	364 854	373 576	382 376	391 255	400 214	409 255	418 378	427 584
0.78	436 875	446 252	455 715	465 267	474 908	484 639	494 462	504 379	514 390	524 496
0.79	534 700	545 002	555 404	565 908	576 514	587 225	598 041	608 965	619 998	631 141
0.80	2.642 397	653 767	665 252	676 855	688 577	700 420	712 386	724 476	736 694	749 040

TABLE II. Values of the integral I_2.

b	0	1	2	3	4	5	6	7	8	9
0.00	0.000 000	000 750	001 500	002 250	003 000	003 750	004 500	005 250	006 000	006 751
0.01	007 501	008 251	009 001	009 752	010 502	011 253	012 003	012 754	013 505	014 256
0.02	015 007	015 758	016 509	017 260	018 011	018 763	019 514	020 266	021 018	021 770
0.03	022 522	023 274	024 027	024 780	025 532	026 285	027 038	027 792	028 545	029 299
0.04	030 053	030 807	031 561	032 315	033 070	033 825	034 580	035 335	036 091	036 847
0.05	0.037 603	038 359	039 116	039 872	040 630	041 387	042 145	042 902	043 661	044 419
0.06	045 178	045 937	046 696	047 456	048 216	048 976	049 737	050 498	051 259	052 021
0.07	052 783	053 545	054 308	055 071	055 834	056 598	057 362	058 127	058 892	059 657
0.08	060 423	061 189	061 955	062 722	063 490	064 258	065 026	065 794	066 564	067 333
0.09	068 103	068 873	069 644	070 416	071 188	071 960	072 733	073 506	074 280	075 054
0.10	0.075 829	076 604	077 380	078 156	078 933	079 711	080 488	081 267	082 046	082 825
0.11	083 606	084 386	085 168	085 949	086 732	087 515	088 298	089 083	089 868	090 653
0.12	091 439	092 226	093 013	093 801	094 589	095 378	096 168	096 959	097 750	098 542
0.13	099 334	100 127	100 921	101 716	102 511	103 307	104 104	104 901	105 699	106 498
0.14	107 297	108 098	108 899	109 700	110 503	111 306	112 110	112 915	113 721	114 527
0.15	0.115 334	116 142	116 951	117 761	118 571	119 382	120 194	121 007	121 821	122 636
0.16	123 451	124 267	125 085	125 903	126 722	127 541	128 362	129 184	130 006	130 830
0.17	131 654	132 479	133 305	134 133	134 961	135 790	136 620	137 451	138 282	139 115
0.18	139 949	140 784	141 620	142 457	143 295	144 134	144 974	145 815	146 657	147 500
0.19	148 344	149 189	150 035	150 883	151 731	152 581	153 431	154 283	155 136	155 990

TABLE II.—*Continued.*

b	0	1	2	3	4	5	6	7	8	9
0.20	0.156 845	157 701	158 558	159 417	160 276	161 137	161 999	162 862	163 726	164 592
0.21	165 459	166 326	167 196	168 066	168 938	169 810	170 684	171 560	172 436	173 314
0.22	174 193	175 074	175 955	176 838	177 723	178 608	179 495	180 383	181 273	182 164
0.23	183 056	183 950	184 845	185 741	186 639	187 538	188 439	189 341	190 244	191 149
0.24	192 055	192 963	193 872	194 783	195 695	196 609	197 524	198 440	199 359	200 278
0.25	0.201 199	202 122	203 046	203 972	204 900	205 828	206 759	207 691	208 625	209 560
0.26	210 497	211 436	212 376	213 318	214 261	215 206	216 153	217 101	218 052	219 004
0.27	219 957	220 912	221 869	222 828	223 789	224 751	225 715	226 681	227 649	228 618
0.28	229 589	230 562	231 537	232 514	233 493	234 473	235 455	236 440	237 426	238 414
0.29	239 404	240 396	241 389	242 385	243 383	244 383	245 384	246 388	247 393	248 401
0.30	0.249 411	250 423	251 436	252 452	253 470	254 490	255 512	256 536	257 563	258 591
0.31	259 622	260 654	261 689	262 726	263 765	264 807	265 851	266 897	267 945	268 995
0.32	270 048	271 103	272 160	273 219	274 281	275 345	276 412	277 481	278 552	279 626
0.33	280 702	281 780	282 861	283 944	285 030	286 118	287 208	288 302	289 397	290 495
0.34	291 596	292 699	293 805	294 913	296 024	297 138	298 254	299 372	300 494	301 618
0.35	0.302 745	303 874	305 006	306 141	307 279	308 419	309 562	310 708	311 856	313 008
0.36	314 162	315 319	316 479	317 642	318 808	319 976	321 148	322 322	323 500	324 680
0.37	325 864	327 050	328 240	329 432	330 628	331 826	333 028	334 233	335 440	336 652
0.38	337 866	339 083	340 304	341 527	342 754	343 985	345 218	346 455	347 695	348 939
0.39	350 185	351 435	352 689	353 946	355 206	356 470	357 737	359 008	360 282	361 560
0.40	0.362 841	364 126	365 414	366 706	368 002	369 301	370 604	371 911	373 221	374 535
0.41	375 852	377 174	378 499	379 828	381 161	382 498	383 839	385 183	386 532	387 884
0.42	389 240	390 601	391 965	393 334	394 706	396 083	397 463	398 848	400 237	401 630
0.43	403 028	404 429	405 835	407 245	408 659	410 078	411 501	412 928	414 360	415 797
0.44	417 237	418 682	420 132	421 586	423 045	424 509	425 977	427 449	428 927	430 409
0.45	0.431 896	433 387	434 883	436 385	437 891	439 401	440 917	442 438	443 963	445 494
0.46	447 030	448 570	450 116	451 667	453 223	454 784	456 351	457 922	459 499	461 082
0.47	462 669	464 262	465 860	467 464	469 074	470 688	472 308	473 934	475 566	477 203
0.48	478 846	480 494	482 148	483 808	485 474	487 146	488 824	490 507	492 197	493 892
0.49	495 594	497 301	499 015	500 735	502 461	504 193	505 932	507 677	509 428	511 186
0.50	0.512 950	514 720	516 498	518 281	520 071	521 868	523 672	525 482	527 299	529 123
0.51	530 954	532 792	534 637	536 488	538 347	540 213	542 086	543 966	545 853	547 748
0.52	549 650	551 559	553 476	555 400	557 332	559 271	561 218	563 173	565 135	567 106
0.53	569 084	571 070	573 064	575 065	577 075	579 093	581 119	583 154	585 197	587 248
0.54	589 307	591 375	593 451	595 536	597 629	599 732	601 842	603 962	606 091	608 228
0.55	0.610 375	612 530	614 695	616 869	619 052	621 244	623 446	625 657	627 878	630 108
0.56	632 348	634 598	636 857	639 126	641 405	643 694	645 994	648 303	650 622	652 952
0.57	655 292	657 643	660 004	662 376	664 758	667 152	669 556	671 970	674 396	676 833
0.58	679 281	681 740	684 211	686 693	689 187	691 692	694 208	696 737	699 277	701 829
0.59	704 393	706 970	709 558	712 159	714 772	717 398	720 036	722 687	725 351	728 027
0.60	0.730 717	733 419	736 135	738 864	741 607	744 362	747 132	749 915	752 712	755 523
0.61	758 348	761 187	764 040	766 908	769 790	772 687	775 598	778 525	781 466	784 422
0.62	787 394	790 381	793 383	796 401	799 434	802 484	805 549	808 631	811 728	814 842
0.63	817 973	821 120	824 284	827 465	830 663	833 878	837 111	840 361	843 628	846 914
0.64	850 217	853 538	856 878	860 236	863 613	867 009	870 423	873 856	877 309	880 781
0.65	0.884 273	887 784	891 315	894 867	898 438	902 031	905 643	909 277	912 931	916 607
0.66	920 304	924 023	927 764	931 526	935 311	939 118	942 948	946 800	950 676	954 574
0.67	958 496	962 442	966 412	970 406	974 424	978 466	982 534	986 626	990 744	994 887
0.68	999 056	*003 251	*007 473	*011 720	*015 995	*020 296	*024 625	*028 982	*033 366	*037 778
0.69	1.042 219	046 688	051 186	055 714	060 270	064 857	069 474	074 121	078 799	083 508
0.70	1.088 249	093 021	097 825	102 661	107 531	112 433	117 368	122 337	127 340	132 378
0.71	137 450	142 558	147 701	152 880	158 095	163 346	168 635	173 962	179 326	184 728
0.72	190 169	195 649	201 169	206 729	212 329	217 970	223 652	229 376	235 142	240 951
0.73	246 804	252 699	258 640	264 624	270 654	276 730	282 852	289 021	295 237	301 501
0.74	307 813	314 175	320 586	327 047	333 559	340 123	346 738	353 406	360 128	366 903
0.75	1.373 733	380 618	387 559	394 556	401 611	408 724	415 896	423 127	430 418	437 770
0.76	445 184	452 661	460 201	467 805	475 474	483 209	491 010	498 879	506 817	514 824
0.77	522 901	531 049	539 270	547 563	555 931	564 373	572 892	581 488	590 162	598 916
0.78	607 749	616 664	625 662	634 744	643 910	653 163	662 503	671 932	681 450	691 060
0.79	700 762	710 558	720 449	730 437	740 523	750 708	760 994	771 383	781 876	792 474
0.80	1.803 179	813 994	824 918	835 955	847 106	858 372	869 756	881 259	892 884	904 631

TABLE III.

b	I_1	I_2	I_1-I_2	$(1-b)I_1$	$(1-b)I_2$
0.800	2.642 397	1.803 179	0.839 218	0.528 479	0.360 636
0.801	653 767	813 994	839 773	528 100	360 985
0.802	665 252	824 918	840 334	527 720	361 334
0.803	676 855	835 955	840 900	527 340	361 683
0.804	688 577	847 106	841 471	526 961	362 033
0.805	2.700 420	1.858 372	0.842 048	0.526 582	0.362 383
0.806	712 386	869 756	842 629	526 203	362 733
0.807	724 476	881 259	843 217	525 824	363 083
0.808	736 694	892 884	843 810	525 445	363 434
0.809	749 040	904 631	844 408	525 067	363 785
0.810	2.761 517	1.916 504	0.845 012	0.524 688	0.364 136
0.811	744 126	928 504	845 622	524 310	364 487
0.812	786 871	940 633	846 238	523 932	364 839
0.813	799 752	952 893	846 859	523 554	365 191
0.814	812 773	965 287	847 486	523 176	365 543
0.815	2.825 936	1.977 816	0.848 119	0.522 798	0.365 896
0.816	839 242	1.990 483	848 758	522 420	366 249
0.817	852 694	2.003 290	849 404	522 043	366 602
0.818	866 295	016 240	850 055	521 666	366 956
0.819	880 047	029 334	850 713	521 288	367 310
0.820	2.893 952	2.042 576	0.851 376	0.520 911	0.367 664
0.821	908 014	055 967	852 047	520 534	368 018
0.822	922 234	069 511	852 723	520 158	368 373
0.823	936 616	083 210	853 406	519 781	368 728
0.824	951 161	097 066	854 096	519 404	369 083
0.825	2.965 874	2.111 082	0.854 792	0.519 028	0.369 439
0.826	980 756	125 262	855 495	518 652	369 796
0.827	2.995 812	139 607	856 204	518 275	370 152
0.828	3.011 043	154 121	856 921	517 899	370 509
0.829	026 452	168 808	857 645	517 523	370 866
0.830	3.042 044	2.183 669	0.858 375	0.517 147	0.371 224
0.831	057 821	198 708	859 113	516 772	371 582
0.832	073 786	213 928	859 858	516 396	371 940
0.833	089 943	229 333	860 610	516 020	372 299
0.834	106 295	244 926	861 369	515 645	372 658
0.835	3.122 846	2.260 710	0.862 136	0.515 270	0.373 017
0.836	139 600	276 689	862 911	514 894	373 377
0.837	156 559	292 867	863 693	514 519	373 737
0.838	173 729	309 246	864 483	514 144	374 098
0.839	191 112	325 832	865 280	513 769	374 459
0.840	3.208 714	2.342 628	0.866 086	0.513 394	0.374 820
0.841	226 537	359 638	866 899	513 019	375 182
0.842	244 586	376 865	867 721	512 645	375 545
0.843	262 866	394 315	868 551	512 270	375 907
0.844	281 381	411 991	869 390	511 895	376 271
0.845	3.300 135	2.429 898	0.870 237	0.511 521	0.376 634
0.846	319 133	448 041	871 092	511 147	376 998
0.847	338 380	466 424	871 956	510 772	377 363
0.848	357 881	485 052	872 829	510 398	377 728
0.849	377 640	503 929	873 711	510 024	378 093
0.850	3.397 664	2.523 062	0.874 602	0.509 650	0.378 459
0.851	417 956	542 454	875 502	509 275	378 826
0.852	438 523	562 112	876 411	508 901	379 193
0.853	459 370	582 041	877 330	508 527	379 560
0.854	480 504	602 246	878 258	508 154	379 928
0.855	3.501 929	2.622 733	0.879 196	0.507 780	0.380 296
0.856	523 652	643 508	880 143	507 406	380 665
0.857	545 679	664 578	881 101	507 032	381 035
0.858	568 016	685 948	882 068	506 658	381 405
0.859	590 671	707 625	883 046	506 285	381 775
0.860	3.613 650	2.729 615	0.884 035	0.505 911	0.382 146
0.861	636 959	751 926	885 033	505 537	382 518
0.862	660 607	774 564	886 043	505 164	382 890
0.863	684 600	797 537	887 063	504 790	383 263
0.864	708 946	820 851	888 094	504 417	383 636
0.865	3.733 653	2.844 516	0.889 137	0.504 043	0.384 010
0.866	758 729	868 538	890 191	503 670	384 384
0.867	784 182	892 925	891 256	503 296	384 759
0.868	810 020	917 687	892 333	502 923	385 135
0.869	836 254	942 832	893 422	502 549	385 511
0.870	3.862 891	2.968 368	0.894 523	0.502 176	0.385 888
0.871	889 940	2.994 304	895 636	501 802	386 265
0.872	917 413	3.020 651	896 762	501 429	386 643
0.873	945 318	047 418	897 900	501 055	387 022
0.874	3.973 666	074 615	899 051	500 682	387 402
0.875	4.002 468	3.102 253	0.900 215	0.500 308	0.387 782
0.876	031 734	130 341	901 393	499 935	388 162
0.877	061 476	158 892	902 583	499 562	388 544
0.878	091 705	187 917	903 788	499 188	388 926
0.879	122 434	217 427	905 007	498 814	389 309

TABLE III.—*Continued.*

b	I_1	I_2	I_1-I_2	$(1-b)I_1$	$(1-b)I_2$
0.880	4.153 674	3.247 435	0.906 239	0.498 441	0.389 692
0.881	185 439	277 953	907 486	498 067	390 076
0.882	217 743	308 995	908 748	497 694	390 461
0.883	250 598	340 573	910 025	497 320	390 847
0.884	284 019	372 703	911 317	496 946	391 234
0.885	4.318 022	3.405 398	0.912 624	0.496 572	0.391 621
0.886	352 620	438 673	913 947	496 199	392 009
0.887	387 830	472 544	915 286	495 825	392 397
0.888	423 668	507 027	916 641	495 451	392 787
0.889	460 152	542 139	918 013	495 077	393 177
0.890	4.497 298	3.577 896	0.919 402	0.494 703	0.393 569
0.891	535 125	614 317	920 808	494 329	393 961
0.892	573 652	651 421	922 232	493 954	394 353
0.893	612 899	689 226	923 673	493 580	394 747
0.894	652 885	727 752	925 133	493 206	395 142
0.895	4.693 631	3.767 020	0.926 611	0.492 831	0.395 537
0.896	735 161	807 053	928 108	492 457	395 933
0.897	777 496	847 871	929 624	492 082	396 331
0.898	820 659	889 499	931 161	491 707	396 729
0.899	864 677	931 960	932 717	491 332	397 128
0.900	4.909 574	3.975 280	0.934 294	0.490 957	0.397 528
0.901	4.955 376	4.019 485	935 891	490 582	397 929
0.902	5.002 112	064 602	937 510	490 207	398 331
0.903	049 810	110 659	939 151	489 832	398 734
0.904	098 501	157 687	940 814	489 456	399 138
0.905	5.148 215	4.205 715	0.942 500	0.489 080	0.399 543
0.906	198 985	254 776	944 209	488 705	399 949
0.907	250 845	304 903	945 942	488 329	400 356
0.908	303 831	356 132	947 699	487 952	400 764
0.909	357 980	408 499	949 481	487 576	401 173
0.910	5.413 330	4.462 042	0.951 288	0.487 200	0.401 584
0.911	469 922	516 801	953 121	486 823	401 995
0.912	527 798	572 817	954 981	486 446	402 408
0.913	587 002	630 134	956 868	486 069	402 822
0.914	647 580	688 798	958 782	485 692	403 237
0.915	5.709 582	4.748 856	0.960 726	0.485 314	0.403 653
0.916	773 057	810 359	962 698	484 937	404 070
0.917	838 059	873 359	964 700	484 559	404 489
0.918	904 644	4.937 911	966 733	484 181	404 909
0.919	5.972 869	5.004 072	968 797	483 802	405 330
0.920	6.042 798	5.071 904	0.970 894	0.483 424	0.405 752
0.921	114 493	141 470	973 023	483 045	406 176
0.922	188 024	212 837	975 187	482 666	406 601
0.923	263 460	286 076	977 385	482 286	407 028
0.924	340 879	361 260	979 618	481 907	407 456
0.925	6.420 358	5.438 469	0.981 889	0.481 527	0.407 885
0.926	501 981	517 783	984 197	481 147	408 316
0.927	585 836	599 292	986 544	480 766	408 748
0.928	672 015	683 085	988 931	480 385	409 182
0.929	760 618	769 260	991 358	480 004	409 617
0.930	6.851 748	5.857 920	0.993 828	0.479 622	0.410 054
0.931	6.945 514	5.949 173	996 341	479 240	410 493
0.932	7.042 032	6.043 133	0.998 899	478 858	410 933
0.933	141 426	139 923	1.001 503	478 476	411 375
0.934	243 827	239 672	004 154	478 093	411 818
0.935	7.349 371	6.342 517	1.006 854	0.477 709	0.412 264
0.936	458 208	448 603	009 605	477 325	412 711
0.937	570 494	558 086	012 408	476 941	413 159
0.938	686 394	671 130	015 264	476 556	413 610
0.939	806 087	787 911	018 176	476 171	414 063
0.940	7.929 762	6.908 617	1.021 146	0.475 786	0.414 517
0.941	8.057 622	7.033 447	024 175	475 400	414 973
0.942	189 882	162 617	027 265	475 013	415 432
0.943	326 774	296 355	030 419	474 626	415 892
0.944	468 546	434 906	033 639	474 239	416 355
0.945	8.615 463	7.578 535	1.036 928	0.473 850	0.416 819
0.946	767 811	727 524	040 287	473 462	417 286
0.947	8.925 898	7.882 177	043 720	473 073	417 755
0.948	9.090 053	8.042 823	047 230	472 683	418 227
0.949	260 634	209 815	050 819	472 292	418 701
0.950	9.438 025	8.383 534	1.054 491	0.471 901	0.419 177
0.951	9.622 644	8.564 395	058 249	471 510	419 655
0.952	9.814 941	8.752 844	062 097	471 117	420 137
0.953	10.015 406	8.949 368	066 038	470 724	420 620
0.954	10.224 571	9.154 494	070 077	470 330	421 107
0.955	10.443 017	9.368 798	1.074 219	0.469 936	0.421 596
0.956	10.671 373	9.592 907	078 467	469 540	422 088
0.957	10.910 333	9.827 507	083 826	469 144	422 583
0.958	11.160 652	10.073 349	087 303	468 747	423 081
0.959	11.423 161	10.331 258	091 903	468 350	423 582

TABLE III.—*Continued.*

b	I_1	I_2	I_1-I_2	$(1-b)I_1$	$(1-b)I_2$
0.960	11.698 773	10.602 141	1.096 632	0.467 951	0.424 086
0.961	11.988 496	10.887 000	101 496	467 551	424 593
0.962	12.293 442	11.186 938	106 503	467 151	425 104
0.963	12.614 845	11.503 184	111 661	466 749	425 618
0.964	12.954 075	11.837 097	116 977	466 347	426 136
0.965	13.312 659	12.190 197	1.122 462	0.465 943	0.426 657
0.966	13.692 304	12.564 179	128 125	465 538	427 182
0.967	14.094 922	12.960 947	133 976	465 132	427 711
0.968	14.522 667	13.382 640	140 027	464 725	428 244
0.969	14.977 968	13.831 675	146 292	464 317	428 782
0.970	15.463 578	14.310 793	1.152 785	0.463 907	0.429 324
0.971	15.982 632	14.823 111	159 520	463 496	429 870
0.972	16.538 710	15.372 193	166 516	463 084	430 421
0.973	17.135 923	15.962 131	173 792	462 670	430 978
0.974	17.779 016	16.597 647	181 369	462 254	431 539
0.975	18.473 491	17.284 220	1.189 270	0.461 837	0.432 106
0.976	19.225 766	18.028 243	197 523	461 418	432 678
0.977	20.043 378	18.837 219	206 159	460 998	433 256
0.978	20.935 231	19.720 020	215 210	460 575	433 840
0.979	21.911 926	20.687 208	224 717	460 150	434 431
0.980	22.986 183	21.751 458	1.234 725	0.459 724	0.435 029
0.981	24.173 400	22.928 116	245 285	459 295	435 634
0.982	25.492 397	24.235 940	256 457	458 863	436 247
0.983	26.966 418	25.698 105	268 313	458 429	436 868
0.984	28.624 519	27.343 584	280 935	457 992	437 497
0.985	30.503 504	29.209 079	1.294 425	0.457 553	0.438 136
0.986	32.650 687	31.341 786	308 901	457 110	438 785
0.987	35.127 941	33.803 428	324 513	456 663	439 445
0.988	38.017 758	36.676 316	341 442	456 213	440 116
0.989	41.432 623	40.072 702	359 921	455 759	440 800
0.990	45.530 008	44.149 762	1.380 246	0.455 300	0.441 498
0.991	50.537 365	49.134 555	402 810	454 836	442 211
0.992	56.795 854	55.367 710	428 144	454 367	442 942
0.993	64.841 565	63.384 573	456 992	453 891	443 692
0.994	75.567 942	74.077 495	490 447	453 408	444 465
0.995	90.583 117	89.052 914	1.530 203	0.452 916	0.445 265
0.996	113.103 214	111.524 117	579 098	452 413	446 096
0.997	150.632 202	148.989 745	642 457	451 897	446 969
0.998	225.680 980	223.948 737	732 243	451 362	447 897
0.999	450.798 638	448.911 987	886 651	450 799	448 912
1.000	∞	∞	∞	0.450 158	0.450 158

TABLE IV.

b	I_1	I_2	I_1-I_2	$(1-b)I_1$	$(1-b)I_2$
0.9950	90.583 12	89.052 91	1.530 203	0.452 9156	0.445 2646
1	92.421 60	90.886 98	534 619	452 8658	445 3462
2	94.336 66	92.797 53	539 129	452 8160	445 4281
3	96.333 19	94.789 45	543 736	452 7660	445 5104
4	98.416 50	96.868 06	548 445	452 7159	445 5931
0.9955	100.592 38	99.039 12	1.553 259	0.452 6657	0.445 6760
6	102.867 13	101.308 95	558 184	452 6154	445 7594
7	105.247 66	103.684 44	563 225	452 5649	445 8431
8	107.741 52	106.173 13	568 386	452 5144	445 9272
9	110.357 00	108.783 32	573 675	452 4637	446 0116
0.9960	113.103 21	111.524 12	1.579 098	0.452 4129	0.446 0965
1	115.990 23	114.405 57	584 660	452 3619	446 1817
2	119.029 16	117.438 79	590 370	452 3108	446 2674
3	122.232 32	120.636 08	596 235	452 2596	446 3535
4	125.613 39	124.011 12	602 264	452 2082	446 4400
0.9965	129.187 62	127.579 15	1.608 466	0.452 1567	0.446 5270
6	132.972 05	131.357 20	614 851	452 1050	446 6145
7	136.985 80	135.364 37	621 430	452 0531	446 7024
8	141.250 35	139.622 14	628 215	452 0011	446 7908
9	145.789 99	144.154 77	635 219	451 9490	446 8798
0.9970	150.632 20	148.989 75	1.642 457	0.451 8966	0.446 9692
1	155.808 30	154.158 36	649 944	451 8441	447 0592
2	161.354 06	159.696 36	657 698	451 7914	447 1498
3	167.310 54	165.644 80	665 739	451 7385	447 2410
4	173.725 13	172.051 04	674 088	451 6853	447 3327
0.9975	180.652 81	178.970 04	1.682 769	0.451 6320	0.447 4251
6	188.157 70	186.465 89	691 809	451 5785	447 5181
7	196.315 10	194.613 86	701 239	451 5247	447 6119
8	205.213 96	203.502 87	711 094	451 4707	447 7063
9	214.960 22	213.238 81	721 414	451 4165	447 8015
0.9980	225.680 98	223.948 74	1.732 243	0.451 3620	0.447 8975
1	237.530 09	235.786 46	743 635	451 3072	447 9943
2	250.695 61	248.939 96	755 649	451 2521	448 0919
3	265.409 84	263.641 48	768 358	451 1967	448 1905
4	281.963 14	280.181 29	781 846	451 1410	448 2901
0.9985	300.723 32	298.927 11	1.796 213	0.451 0850	0.448 3907
6	322.163 26	320.351 68	811 581	451 0286	448 4924
7	346.901 35	345.073 26	828 099	450 9718	448 5952
8	375.762 11	373.916 16	845 950	450 9145	448 6994
9	409.869 85	408.004 48	865 368	450 8568	448 8049
0.9990	450.798 6	448.912 0	1.886 651	0.450 7986	0.448 9120
1	500.822 1	498.911 9	910 194	450 7399	449 0207
2	563.350 6	561.414 1	936 530	450 6805	449 1313
3	643.743 5	641.777 1	1.966 408	450 6204	449 2440
4	750.932 6	748.931 7	2.000 923	450 5596	449 3590
0.9995	900.995 6	898.953 8	2.041 775	0.450 4978	0.449 4769
6	1126.087 2	1123.995 3	091 809	450 4349	449 5981
7	1501.235 1	1499.078 7	156 364	450 3705	449 7236
8	2251.521 5	2249.274 0	247 420	450 3043	449 8548
9	4502.351 2	4499.948 0	403 215	450 2351	449 9948
1.0000	∞	∞	∞	0.450 1582	0.450 1582

Effect of Coating Composition of Oxide-Coated Cathodes on Electron Emission

E. G. WIDELL AND R. A. HELLAR
Tube Department, RCA Victor Division, Harrison, New Jersey
(Received June 6, 1950)

Investigation of oxide-coated cathodes has been made to determine the effect on electron emission of varying the proportions of emitting oxides. The results show that maximum electron emission under saturation conditions is obtained from a solid solution containing strontium oxide (SrO) and barium oxide (BaO) in an approximate molecular ratio of 7 to 3.

These results also indicate that maximum size of the co-precipitated alkaline-earth carbonate particle occurs at the same molecular ratio as that giving maximum electron emission from the oxides.

When the saturation emission obtained with a square-wave voltage pulse was measured, the current pulse as observed on a synchroscope was also a square wave and showed no measurable decay characteristic for a pulse duration time of 10 microseconds.

INTRODUCTION

MANY investigations have been made of the thermionic emission from alkaline-earth oxide-coated cathodes. However, only a few have considered the relation between emission capability and coating composition. In general, coatings have followed a 50–50 ratio of barium and strontium oxides either as weight percent or as molecular percent and have been applied in the form of a mechanical mixture or co-precipitate of the carbonates.

Arnold[1] in 1920 reported that the highest emission was obtained from BaO alone but an equal quantity of SrO was added to improve life. Simon[2] in 1927 found two emission maxima, the first at about 85 percent SrO–15 percent BaO, and the second at about 85 percent BaO–15 percent SrO (weight percent). BaO alone yielded about 70 percent of the emission obtained from a 50–50 ratio. No other investigator finds two maxima. In 1933, Benjamin and Rooksby[3] reported maximum electron emission from a molecular weight ratio of 63 percent SrO and 37 percent BaO. Huber[4] in 1941 obtained a maximum for the BaO–SrO system at about 55 percent BaO with an emission value about ten times as great as from BaO alone. Veenemans[5] in 1943 published a curve showing maximum emission at approximately 75 molecular percent of SrO. Weinreich[6] reported in 1945 that a BaO–SrO cathode coating containing 98.2 percent SrO yielded approximately the same surface emission as coatings containing lesser amounts of SrO. Eisenstein[7] in 1946 stated that Fineman of the M.I.T. Radiation Laboratory obtained the same pulsed electron emission from BaO alone as from a coating containing 50 percent BaO–50 percent SrO.

In view of all these apparent contradictions and because present-day requirements for high electron-emission currents under pulsed conditions again focus attention on types of oxide coatings which yield maximum currents, it was thought desirable to reinvestigate this subject.

MATERIALS AND CONSTRUCTION

The carbonates were prepared by co-precipitation from nitrate solutions using sodium carbonate as the precipitating agent with constant conditions of precipitation except for change in ratios of the two components. The strontium carbonate actually contained 1.0 percent $BaCO_3$ while the barium carbonate contained 0.2 percent $SrCO_3$, but for purposes of comparison the values are listed as 100 percent for both carbonates and oxides.

The cathode sleeves were made of Grade *A* nickel which, by spectrographic analysis, contained 0.2 percent silicon, 0.045 percent magnesium, 0.10 percent manganese, 0.07 percent iron, 0.04 percent copper, and less than 0.02 percent titanium. Oxide inclusions were not measured. The sleeves were 25 millimeters long and had an outside diameter of 0.762 millimeters. They were cleaned by electropolishing before being coated in order to remove surface contaminations.

The carbonates were milled for 45 hours under identical conditions in the usual organic solvents with nitrocellulose as a binder, and were sprayed on the central 10-millimeter portion of the cathode sleeves. The average weight of the sprayed coating was 3.4 mg/cm^2 with a thickness of 0.044 millimeter giving an average coating density of about 0.73 g/cm^3.

At least five tubes were used for testing each ratio of oxides in order to obtain more dependable results than could be obtained by the use of a single tube. In order to obtain even temperatures throughout the length of the coated cathodes, only 10 millimeters of the central portions of each cathode sleeve were coated. Optical pyrometer measurements showed a variation of less than 10°K between the center and the ends of the coated portion. The bare ends of the sleeves were carefully cleaned to prevent unwanted electron emission. Tests with a specially constructed tube showed that the coated portion gave a thousand times more emission than the bare portions; therefore, emission from the

[1] H. D. Arnold, Phys. Rev. **16**, 70 (1920).
[2] H. Simon, Zeits. f. techn. Phys. **8**, 434 (1927).
[3] M. Benjamin and H. P. Rooksby, Phil. Mag. **15**, 810 (1933).
[4] H. Huber, dissertation, Berlin, 1941.
[5] C. P. Veenemans, Nederl. Tijdschr. Natuurkunde **10**, 1 (1943).
[6] M. O. Weinreich, Rev. Gén. Elec. (August, 1945).
[7] A. Eisenstein, J. App. Phys. **17**, 654 (1946).

TABLE I.

Carbonates					
Molecular % $BaCO_3$	Molecular % $SrCO_3$	"Average" particle size, microns	Bulk density, g/cm³	Sprayed cathode-coating density, g/cm³	Cathode brightness temperature, °K
0	100	1.32	0.46	0.76	998
10	90	1.87	0.39	0.85	1010
20	80	3.08	0.35	0.73	1000
30	70	3.35	0.33	0.67	996
40	60	2.92	0.37	0.73	998
50	50	2.83	0.40	0.73	1006
60	40	2.59	0.42	0.74	989
70	30	2.11	—	0.74	991
80	20	1.96	0.48	0.71	992
100	0	1.23	0.63	0.79	975

bare portions could be neglected. The cathodes were mounted in a standard diode in which carbonized plates with inside diameters of 1.78 millimeters were used. Each tube received the same exhaust on an all-glass system using liquid air with mercury diffusion pumps and was sealed off after flashing a Ba—Mg getter. All tubes were given the same aging schedule.

MEASUREMENTS

Structural Characteristics of the Carbonates

The particle sizes of the carbonates as precipitated were measured by the Andreasen pipette sedimentation method. The "average" particle size reported is the particle size for the 50 percent point on the curve showing cumulative weight percent *versus* particle size. These data are shown in Table I and are compared with emission data in Fig. 6. X-ray diffraction patterns indicate that the carbonate crystals of the 50–50 ratio have an aragonite structure. Electron microscope pictures of 100 percent $BaCO_3$, 70 percent $BaCO_3$–30 percent $SrCO_3$, 30 percent $BaCO_3$–70 percent $SrCO_3$, and 100 percent $SrCO_3$ are shown in Fig. 1. These pictures show the change in particle size with change in ratios of the carbonates. Particle or crystallite measurements of the 70 percent $BaCO_3$–30 percent $SrCO_3$ were made by the method of low angle scattering of x-rays using $K\alpha$-radiation at 35 kilovolts and 15 milliamperes, filtered by nickel; the x-ray film was spaced 50 centimeters from the specimens. These results are shown in Table II. The particle size as determined by this method may be the measurements of single or small groups of crystallites since the diameters are less than 200 angstroms, whereas the particle sizes as measured by the electron microscope or by the Andreasen pipette are a hundred times greater. The above data show that both the particles and the crystallites of the 70 percent $SrCO_3$–30 percent $BaCO_3$ are about 1.5 times larger than those of the 30 percent $SrCO_3$–70 percent $BaCO_3$.

Cathode Temperatures

Cathode temperature measurements were made with a disappearing-filament type of optical pyrometer ($\lambda = 0.665\mu$). The cathodes were initially adjusted to a brightness temperature of about 1000°K in order to

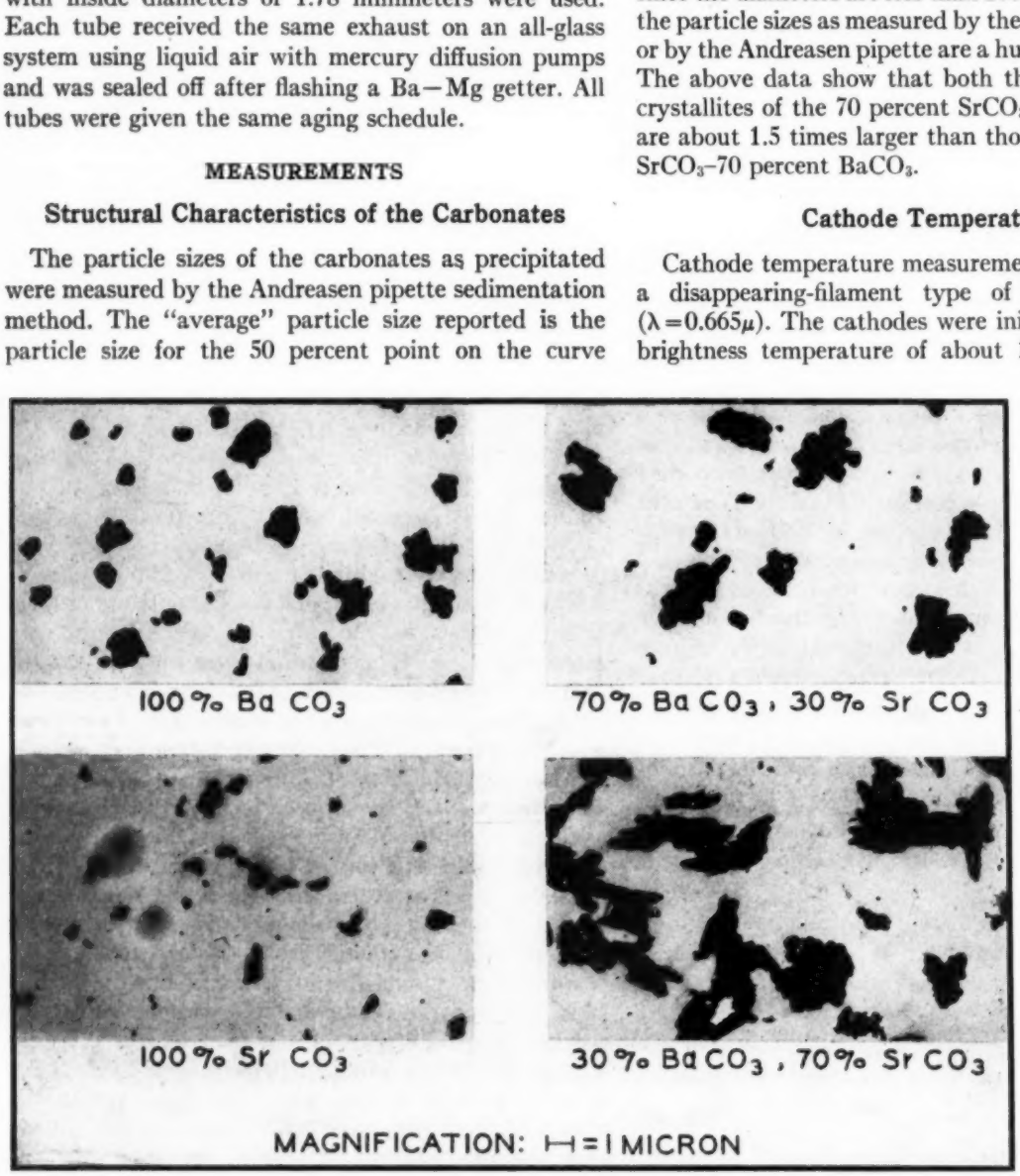


FIG. 1. Electron micrographs of carbonate particles.

TABLE II. Particle size measurement by low angle scattering of x-rays.*

Type of carbonate	$r=1.0$	1.5	2.0	2.5	3.0
70% $SrCO_3$–30% $BaCO_3$	$R=105$	79	67	48	38
30% $SrCO_3$–70% $BaCO_3$	$R=63$	58	44	34	32

* r = distance from center of beam on film in millimeters. R = radius of particle in angstrom units.

limit electron emission and thus limit the heating up of the tube parts and cathode coatings during testing. The true temperature was then about 1030°K based on a spectral-emissivity (0.665μ) value of 0.56 for this type of coating obtained from unpublished data in this laboratory. The cathode temperatures shown in Table I are brightness temperatures (with no correction for glass absorption) and are averages for five tubes in each test.

Electron Emission

The emission measurements were made on a pulse emission tester designed to produce microsecond rectangular pulses. The pulse duration was set at 3 microseconds and the repetition rate at 500 pulses per second. A synchroscope showed wave forms of voltage and current to be square. Each measurement plotted represents an averaged reading for five or more tubes. The results are shown in Figs. 2, 3, and 4. Figure 5 compares these data with those of Benjamin and Rooksby, and of Veenemans.

Emission Decay

For measurements of relative values of emission for the various ratios of oxides under pulse-testing conditions, it may be necessary to consider the effect of emission decay characteristics as reported by Sproull[8] who found a considerable decay of emission during a 10-microsecond pulse. Coomes[9] also mentions this effect, stating that Ramsay of the Bartol Foundation found

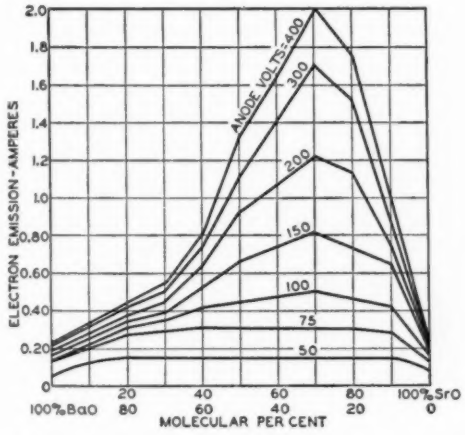


FIG. 2. Electron emission variation with coating compositions at different anode potentials.

[8] R. L. Sproull, Phys. Rev. **67**, 166 (1945).
[9] E. A. Coomes, J. App. Phys. **17**, 647 (1946).

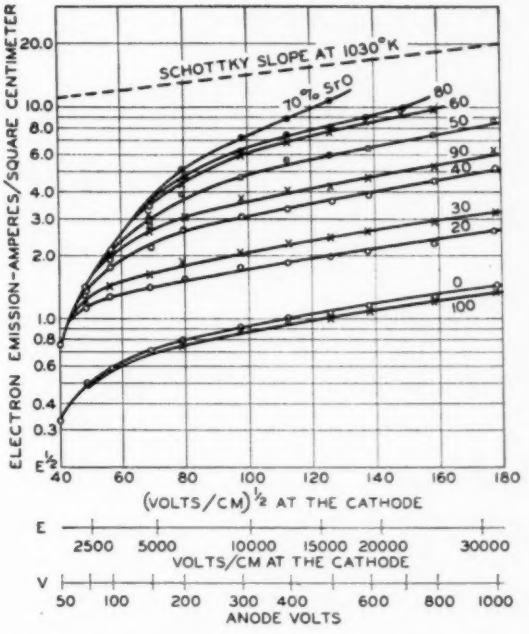


FIG. 3. Electron emission for various coating compositions. Schottky plot. $\mathrm{Log}_{10}(I/\mathrm{cm}^2)$ *versus* $E^{\frac{1}{2}}$, E, V.

short-time (10^{-5} second) emission decay in aged cathodes but not in very active cathodes. Observations for this effect were made on the cathodes with a 10-microsecond pulse. The emission pulse, observed on a synchroscope, had a square wave form and showed no measurable decay during this pulse period.

Decay effects have been observed in this laboratory from cathodes in tubes of the same construction but made for testing different materials. The decay effect in this case, about 40 percent decay during a 10-microsecond pulse, occurred only in those tubes which showed a heavy cathode poisoning effect as evidenced by

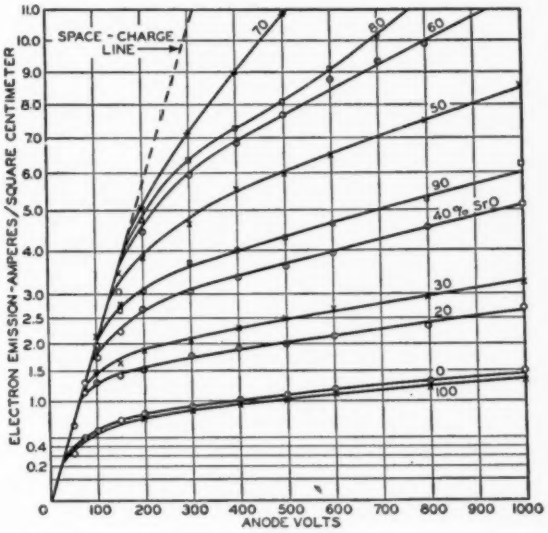


FIG. 4. Electron emission for various coating compositions. Space charge plot. $(I/\mathrm{cm}^2)^{\frac{1}{2}}$ *versus* anode volts.

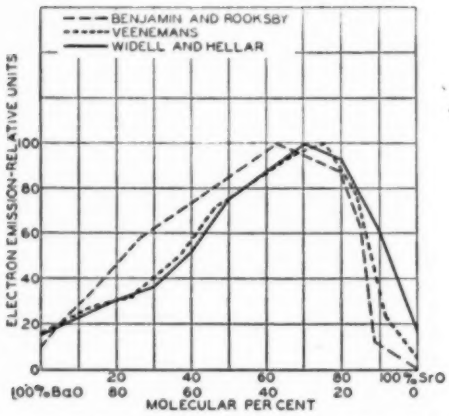


FIG. 5. Comparison of the results of different investigators.

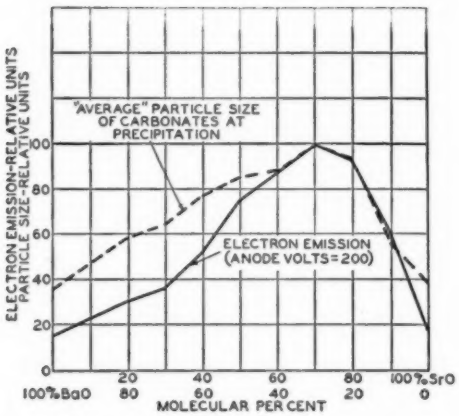


FIG. 6. Comparison of carbonate particle size and electron emission with coating composition.

slumping emission when measured under steady d.c. testing conditions.

DISCUSSION

The Effect of Oxide Ratios on Electron Emission

Figure 2 shows the variation in electron emission for various percentages of BaO and SrO and at various anode voltages. At saturation the curves peak at 70 percent SrO–30 percent BaO (molecular percent). The emission for 100 percent BaO is lower than that observed in previous tests on similar cathodes but this difference may partly be accounted for by the 25° lower operating temperature of this set as shown in Table I. A previous test by Wolfgang Panofsky* of this laboratory showed a much flatter curve with the emission from the BaO about 80 percent of that from the 50–50 ratio but the emission values for the higher percentages of BaO were not stable and decreased during operation.

Figure 4 shows the emission data plotted according to the $\frac{3}{2}$-power law. The large deviations from the normal in the Schottky region for the very highest emission values are due to heating up of the cathodes which could have been lessened if a shorter pulse duration and a lower pulse repetition rate were used.

Comparison with Previous Investigations

When the data obtained in this investigation are compared with those of previous investigators, as shown in Fig. 5, emission values were selected at an anode voltage of 200 volts (about 6500 volts/cm at the cathode) because at that voltage all of the cathodes approached emission saturation. A better comparison might have been made if all investigators had reported zero field emission. Values of emission are given in terms of percent of maximum to facilitate comparison with the published curves of Benjamin-Rooksby and of Veenemans. All three curves have approximately the same general shape except that the Benjamin-Rooksby curve peaks at 63 percent SrO–37 percent BaO while the curve from this investigation peaks at 70 percent $SrCO_3$–30 percent $BaCO_3$.

The Effect of Particle Size on Electron Emission

There has been some discussion in the literature as to the effect of crystal size and particle size on electron emission from oxide-coated cathodes. The change in particle size with variation in ratios of the carbonates when precipitated under identical conditions was measured and the data are shown in comparison with electron emission data in Fig. 6. The curve for the particle size of the carbonates peaks at the same molecular ratio which yields maximum electron emission from the oxide. Although the agreement between the two curves may be merely a coincidence, this subject is being further investigated.

CONCLUSIONS

Maximum electron emission from oxide-coated cathodes consisting of solid solutions of BaO and SrO occurs when the oxides have a molecular percentage of 70 percent SrO and 30 percent BaO.

The maximum size of the precipitated carbonate crystal in the system of $BaCO_3-SrCO_3$ occurs at the same molecular ratio which gives the maximum electron emission from the oxides.

There is a close relationship between the size of the carbonate particles and the size of their crystallites as shown by comparing the measurements obtained by the Andreasen pipette with those obtained by the low angle scattering of x-rays on the 70–30 and the 30–70 ratios of carbonates.

The tubes of this test exhibited no measurable electron-emission decay during a 10-microsecond pulse duration.

ACKNOWLEDGMENTS

The authors wish to acknowledge their indebtedness to Miss Irene Kakascik for making the electron micrographs, and to Mr. R. B. DeLong for making the determinations of crystal size by low angle scattering of x-rays.

* Now at California Institute of Technology, Pasadena, California.

Surface Waves and Their Application to Transmission Lines

GEORG GOUBAU
Radio Communication Branch, Coles Signal Laboratory, Fort Monmouth, New Jersey
(Received March 10, 1950)

In this paper the applicability of non-radiating surface waves for transmission lines is investigated. Two types of waves are considered. The first one, originally studied by A. Sommerfeld, is guided by a cylindrical conductor of finite conductivity. Although this wave type has (under comparable conditions) much lower attenuation than the waves in coaxial cables or rigid wave guides, its practical application is restricted by the fact that the extension of the field is very large. Efficient excitation and undisturbed propagation of this wave mode are feasible only for very high frequencies. The other wave type considered in this paper has not been treated in the literature. It is guided by a conductor which is coated with a dielectric layer or the surface of which is otherwise modified; for example, by being threaded. The field of this wave type has a structure similar to that of Sommerfeld's wave, but the extension of the field can be controlled by the surface modification. Thus low loss transmission lines on the basis of this wave type become feasible for frequencies above 100 megacycles. The information necessary for the design of such lines is given and the agreement between the theoretically expected transmission losses and the measured transmission losses is checked.

I. INTRODUCTION

IN 1899, A. Sommerfeld[1] published a paper about wave propagation along an infinitely long cylindrical wire of finite conductivity. The type of wave he investigated is a non-radiating mode which has had little consideration in the modern literature, although the attenuation is theoretically much smaller than that of waves in coaxial cables. There may have been some doubt about the possibility of generating this wave type, as there is in the case of Zenneck's[2] ground wave in wave propagation along the earth. The latter wave is guided by a plane interface separating a non-conducting medium from a conducting one, while in Sommerfeld's case the interface is cylindrical. Both waves are possible solutions of Maxwell's equations satisfying given boundary conditions. However, the solutions are special ones insofar as they refer to *plane* electro-magnetic waves and thus provide that the power sources are infinitely far removed. There has been much discussion about the physical reality of Zenneck's ground wave and there is no convincing answer as yet. However, recent experiments on wave propagation along cylindrical conductors prove that cylindrical surface waves can be generated with high efficiency, and thus they can be employed for guided power transmission. The wave type actually used in these experiments differed from Sommerfeld's wire wave insofar as the surface of the conductor was threaded or coated with a dielectric layer. Such modifications of the conductor surface were made in order to reduce the radial extension of the field and hence the dimensions of the excitation device. The field structure of the modified wave is similar to that of Sommerfeld's wave. There is, however, one interesting difference. Sommerfeld's wave exists only for a conductor of finite conductivity. If the conductivity were increased without limit, the extension of the field would increase in such a manner that the power carried by a wave of finite amplitude would become infinite. The wave propagated by a conductor with a thin coat of dielectric or with an otherwise modified surface however, would exist also in the case of perfect conductivity. The field is only slightly affected by the conductivity, provided the conductivity is above a certain limit.

In order to eliminate misunderstandings, it shall be mentioned that the surface waves considered here are not related to the usual waves on long wire antennas, which are extensively discussed in the literature.[3]

The subject of this paper is a discussion of cylindrical surface waves with regard to their practical application to transmission lines. Two kinds of waves are investigated: The waves on a conductor having a plain surface (Sommerfeld's wave) and the waves on a conductor having a modified surface, in particular on a conductor with a dielectric coat.

II. SOMMERFELD'S WIRE WAVE

An infinitely long cylindrical conductor of circular cross section is embedded in an infinite homogeneous dielectric. The particular solution of this boundary-value problem given by Sommerfeld,[1] describes a radially symmetric transverse magnetic wave travelling along the cylinder. There are also solutions describing waves with cyclic periodicity around the cylinder. These solutions have been investigated by D. Hondros.[4] Because most of the field energy of these unsymmetrical modes is inside the conductor their attenuation is extremely large. Thus, they are of no practical interest and will be excluded from further discussion.

If the coordinate system shown in Fig. 1 is used, the field components of the symmetric wave travelling in

[1] A. Sommerfeld, Ann. Phys. u. Chemie (Neue Folge) **67–1**, 233 (1899).
[2] J. Zenneck, Ann. d. Phys. **23**, 846 (1907).
Sufficient information about references 1 and 2 is given in J. A. Stratton, *Electromagnetic Theory*, (McGraw-Hill Book Company, Inc., New York, 1941).
[3] S. Schelkunoff, *Electromagnetic Waves*, (D. Van Nostrand Company, Inc., New York, 1943).
[4] D. Hondros, Ann. d. Phys. **30**, 905 (1909).

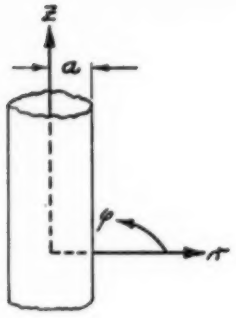


FIG. 1. Coordinate system.

the positive z-direction can be written in the form:

$$E_r = jA\frac{h}{\gamma}Z_1(\gamma r)e^{j(\omega t-hz)},$$

$$E_z = AZ_0(\gamma r)e^{j(\omega t-hz)}, \tag{1}$$

$$H_\varphi = jA\frac{k^2}{\omega\mu\gamma}Z_1(\gamma r)e^{j(\omega t-hz)},$$

as well for the space inside as for the space outside the conductor. The propagation constant h of the guided wave is the same for both parts of the field. The free wave propagation constant k is

$$\text{inside: } k_c = (\omega\mu_c\sigma_c)^{\frac{1}{2}}e^{-j\pi/4}; \quad \text{outside: } k = \omega(\epsilon\mu)^{\frac{1}{2}} \tag{2}$$

(ϵ = dielectric constant, μ = permeability, σ = conductivity). The subscript c refers to the interior of the conductor. The parameters γ and γ_c are defined by

$$\gamma_c^2 = k_c^2 - h^2; \quad \gamma^2 = k^2 - h^2. \tag{3}$$

The cylinder functions Z_0 and Z_1 which describe the field inside the conductor are the Bessel functions J_0 and J_1. They are finite at $r=0$. For the outside the Hankel functions $H_0^{(1)}$ and $H_1^{(1)}$ are taken, since they ensure the proper decrease of the field at infinity and give the proper direction of energy flow if the complex parameter γ is in the first quadrant of the complex plane.

The boundary conditions require that E_z and H_φ be continuous at the surface ($r=a$) of the conductor. Thus, we obtain the equations:

$$\mu\frac{\gamma}{k^2}\frac{H_0^{(1)}(\gamma a)}{H_1^{(1)}(\gamma a)} = \mu_c\frac{\gamma_c}{k_c^2}\frac{J_0(\gamma_c a)}{J_1(\gamma_c a)}, \tag{4}$$

$$\gamma^2 - \gamma_c^2 = k^2 - k_c^2 = k^2\left(1 + j\frac{\sigma\mu_c}{\omega\epsilon\mu}\right) \simeq k^2 j\frac{\sigma\mu_c}{\omega\epsilon\mu}, \tag{5}$$

from which γ and γ_c can be determined and consequently also h. If we restrict our considerations to conductors the radii of which are large compared with the skin depth ($\gamma_c a \gg 1$), J_0 and J_1, in (4) can be replaced by their asymptotic representations:

$$J_0(x) \underset{x\to\infty}{\simeq} \left(\frac{\pi}{2}x\right)^{-\frac{1}{2}}\cos(x-\pi/4);$$

$$J_1(x) \underset{x\to\infty}{\simeq} \left(\frac{\pi}{2}x\right)^{-\frac{1}{2}}\cos(x-3\pi/4). \tag{6}$$

If furthermore the conductor radius is not too large, so that $\gamma a \ll 1$, we can use in (4) the approximations:

$$H_0^{(1)}(x) \underset{x\to 0}{\simeq} \frac{2j}{\pi}\ln(-j0.89x); \quad H_1^{(1)}(x) \underset{x\to 0}{\simeq} -\frac{2j}{\pi x} \tag{7}$$

($0.89\cdots = \frac{1}{2}\times 1.781\cdots = e^C$, where C = Euler's constant $= 0.577\cdots$). Equation (4) is then simplified to

$$-\mu\frac{\gamma^2 a}{k^2}\ln(-j0.89\gamma a) = \mu_c\frac{\gamma_c}{k_c^2}\cot(\gamma_c a - \pi/4) \tag{8}$$

where $\cot(\gamma_c a - \pi/4)$ approaches $+j$ since γ_c becomes complex. We shall see later that all these approximations are justified for the considered conductor diameters and frequency range.

Now, (8) can be written in the form:

$$\xi\ln\xi = \eta \tag{9}$$

with

$$\xi = (-j0.89\gamma a)^2,$$

$$\eta = 2(0.89)^2\frac{k^2 a\mu_c}{|k_c|\mu}e^{j3\pi/4} = |\eta|e^{j3\pi/4}. \tag{10}$$

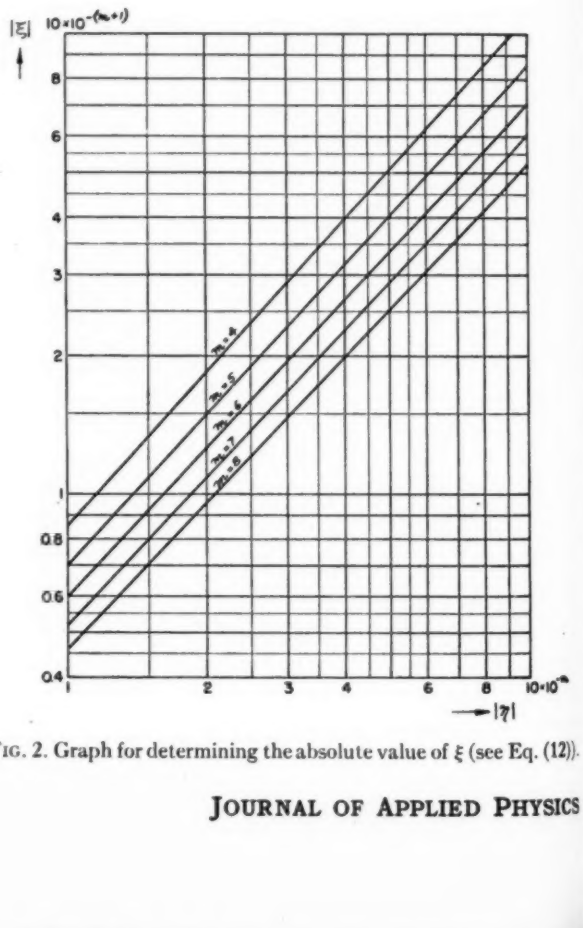


FIG. 2. Graph for determining the absolute value of ξ (see Eq. (12)).

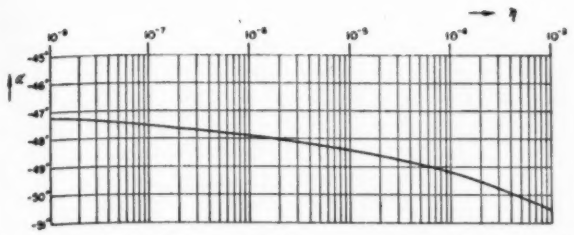


FIG. 3. Graph for determining the argument of ξ (see Eq. (12)).

For copper wire in air

$$|\eta| = 3.27\times10^{-19} a f^{\frac{1}{2}} = 1.70\times10^{-3} a\lambda^{-\frac{1}{2}}. \quad (11)$$

f is the frequency and λ the free space wave-length. λ and a are measured in cm. Equation (9) has been used by Sommerfeld for numerical evaluation of γ. With regard to the following discussion about the applicability of these waves for transmission lines, we are concerned with solutions of (9), within a wide range of f and a. If we consider frequencies from 10^8 to 10^{10} cycles/sec. and radii from 0.1 to 1 cm; then, in the case of a copper wire, $|\eta|$ ranges between 3×10^{-8} and 3×10^{-3}. Within this wide range the complex Eq. (9) can be split into two real equations which are more convenient for numerical evaluation. As shown in the Appendix, the absolute value $|\xi|$ and the argument α of $\xi = |\xi| e^{j\alpha}$ can be determined from the equations:

$$|\xi|\ln|\xi| = -|\eta|; \quad \alpha = -\frac{\pi}{4}\left(1-\frac{1}{\ln|\xi|+1}\right) \quad (12)$$

$|\xi|$ and α are plotted in Figs. 2 and 3 as functions of $|\eta|$. Each of the 5 curves for $|\xi|$ in Fig. 2 refers to one decade in $|\eta|$, identified by the exponent m. α varies only slightly with $|\eta|$ as seen in Fig. 3.

Having $\xi = |\xi| e^{j\alpha}$, one can determine γ from (10):

$$a\gamma = 1.12(|\xi|)^{\frac{1}{2}} e^{j((\alpha+\pi)/2)}. \quad (13)$$

Within the range of $|\eta|$ considered in Figs. 2 and 3 $|a\gamma| < 1.1\times10^{-2}$, $Re[a\gamma_c] > 15$ and $Im[a\gamma_c] > 5$. Hence the simplifications which lead to Eq. (9) are justified.

The propagation constant h of the surface wave is determined by k and γ (see (3)). As γ/k is very small we can write $h \simeq k - \gamma^2/2k$ or, if γ^2 is expressed in terms of ξ with Eq. (10):

$$h = \left(k + 0.63\frac{|\xi|\cos\alpha}{ka^2}\right) + j\left(0.63\frac{|\xi|\sin\alpha}{ka^2}\right). \quad (14)$$

The real part of h yields the phase velocity of the wave:

$$v = c\left(1 - \frac{0.63|\xi|\cos\alpha}{(ka)^2}\right). \quad (15)$$

The phase velocity of the surface wave is somewhat smaller than the phase velocity c of a free wave in the same dielectric. Consider, for instance, a wave with a frequency of 3000 mc, propagating on a copper wire of 0.1 cm radius, stretched in air. From Eq. (11) we obtain $|\eta| = 5.38\times10^{-6}$ and from Figs. 2 and 3 the corresponding $|\xi| = 3.6\times10^{-7}$ and $\alpha = -48°\,15'$. Inserting in (16) we find that the reduction in phase velocity is less than 0.004 percent.

The imaginary part of h gives the attenuation β of the wave. β in nepers per cm is

$$\beta = -\frac{0.63|\xi|\sin\alpha}{ka^2}[n/\text{cm}]. \quad (16)$$

The loss L in db per 100 feet becomes:

$$L = -2.66\frac{\lambda}{a^2}|\xi|\sin\alpha\times10^3[db/100\text{ ft.}]. \quad (17)$$

It may be mentioned that β and L are positive because $\sin\alpha$ is negative.

In Fig. 4, L/100 ft. is plotted for various conductor diameters and a frequency range, $10^2\cdots10^4$ mc ($\lambda = 3\cdots300$ cm). The conductor is assumed to be of copper. It is of interest that even for a wire of only 1 mm radius the losses are considerably smaller than for recommended wave guides. From this fact the application of single conductors for transmission lines appears very promising. However, there is an important question of how to excite efficiently the surface waves. The efficiency of any launching device will be greater the better the agreement between the field built up by the launching device and the field of the wave. The launching device can be considered as a field transformer which transforms the field of a coaxial cable or wave guide into the field of the surface wave. Because the field of the wave extends to infinity, the matching between the field of the launching device and the field of the wave can never be perfect. The larger the launching device the better the chances for good matching conditions. In Section VI we will estimate the loss of a launching device and for this reason we determine in the following that

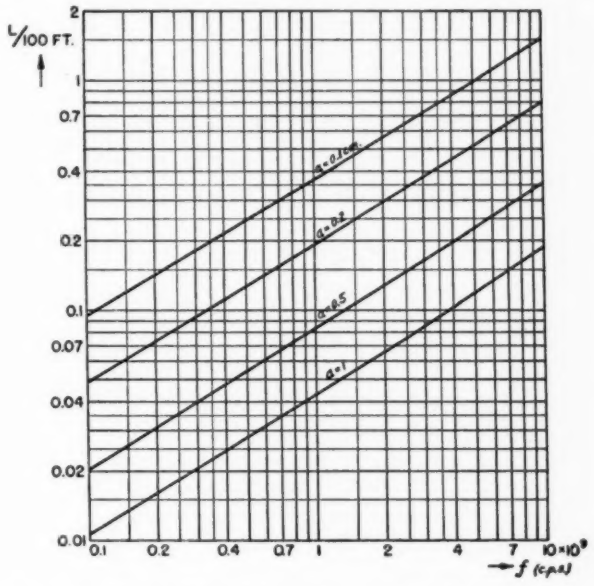


FIG. 4. Transmission loss of copper wires of 0.1, 0.2, 0.5, and 1 cm radius in the case of Sommerfeld's wave.

fraction of the total power of the wave which travels outside of a cylinder of radius ρ surrounding the wire.

The energy flow through the area $\infty > r > \rho$ is given by

$$N_\rho = Re\left[2\pi \int_\rho^\infty r E_r H_\varphi^* dr\right]. \tag{18}$$

(The asterisk denotes the conjugate complex quantity.) The total energy flow N outside the wire is obtained by extending the lower limit of the integral to the wire surface. The quantity we are interested in is N_ρ/N. Inserting for E_r and H_φ the values of Eq. (1) with $Z_0 = H_0^{(1)}$, $Z_1 = H_1^{(1)}$ and considering that $[H^{(1)}(\gamma r)]^* = H^{(2)}(\gamma^* r)$, the power N_ρ at $z=0$ becomes:

$$N_\rho = -Re\left\{2\pi A A^*(\epsilon/\mu)^{\frac{1}{2}} \frac{hk}{\gamma\gamma^*(\gamma^2-\gamma^{*2})} \times[\gamma^*\rho H_1^{(1)}(\gamma\rho) H_0^{(2)}(\gamma^*\rho) - \gamma\rho H_1^{(2)}(\gamma^*\rho) H_0^{(1)}(\gamma\rho)]\right\}. \tag{19}$$

Unfortunately the numerical evaluation of Hankel-functions with complex arguments is cumbersome. However, as we are not too much concerned with high accuracy, we will use the zero representation of the functions up to $|\gamma\rho| < 0.3$, though we have to expect errors of the order of 15 percent for the upper limit of $|\gamma\rho|$. For the wire radii and frequencies considered in Fig. 4, the range $|\gamma\rho| < 0.3$ covers an area around the wire in which the major part of the energy is transmitted.

Using the zero representations of the Hankel functions, shown in (7), and replacing γ by ξ from (10), Eq. (19) becomes:

$$N_\rho = -AA^*(\epsilon/\mu)^{\frac{1}{2}} \frac{ka^4 Re[h]}{|\xi|^2\pi} \times 4(0.89)^4(\ln|\xi|\,\rho^2/a^2 + \alpha\cot\alpha). \tag{20}$$

The total power N is obtained from (20) by setting $\rho = a$. The expression for N becomes very accurate, since $\gamma a \ll 1$. For the ratio N_ρ/N we get:

$$\frac{N_\rho}{N} \simeq 1 + \frac{2\ln(\rho/a)}{\ln|\xi| + \alpha\cot\alpha}. \tag{21}$$

Within the entire range of $|\xi|$ considered above, α is close to $-\pi/4$. Thus

$$\alpha\cot\alpha \simeq (\pi/4) (\ll |\ln|\xi||)$$

and

$$\frac{N_\rho}{N} \simeq 1 + \frac{2\log(\rho/a)}{\log|\xi| + 0.34} \simeq 1 + \frac{2\log(\rho/a)}{\log 2.2|\xi|}. \tag{22}$$

This equation can be used for determining the radius of that area through which a certain percentage of the power is transmitted. Figure 5 shows the half power radius $\rho_{50\%}$ and the radius for 75 percent power $\rho_{75\%}$ within the same frequency range and for the same wire radii as in Fig. 4. The 75 percent radii are not very accurate for reasons mentioned above. Figure 5 shows that the extension of the field is very large, especially for low frequencies and large wire diameters. For instance, for $f = 1000$ mc/sec., $a = 1$ cm, the half power radius is 28 cm and the 75 percent power radius about 150 cm. A transmission line of this dimension would require a clearance around the conductor of several meters, if frequencies of 1000 mc or less should be transmitted.

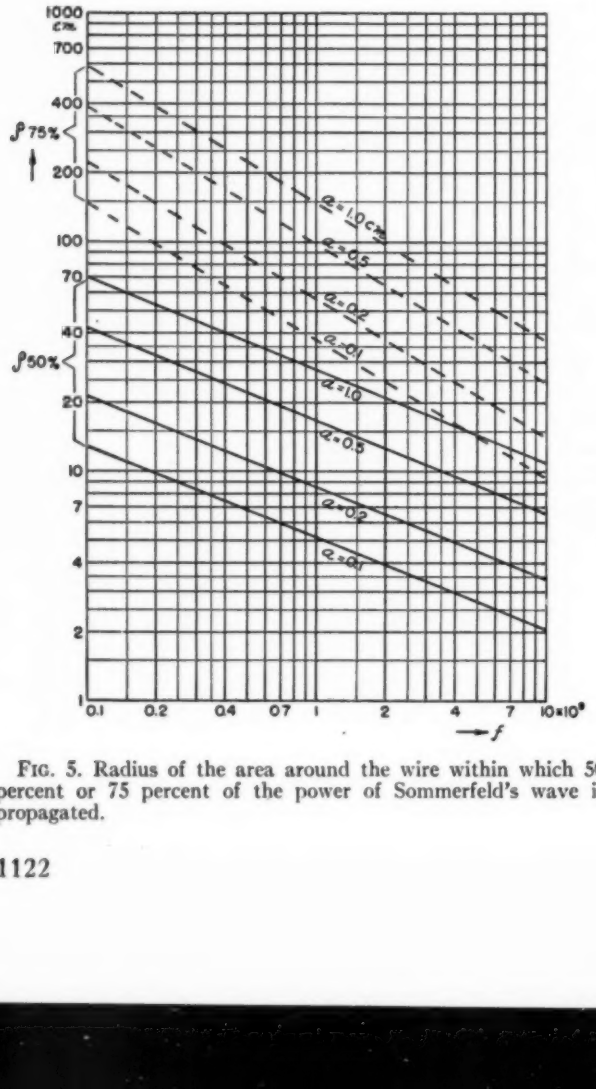


FIG. 5. Radius of the area around the wire within which 50 percent or 75 percent of the power of Sommerfeld's wave is propagated.

In the introduction it was stated that the existence of Sommerfeld's surface wave requires a conductor of finite conductivity. If σ were increased to infinity, $|\eta|$ would become zero and consequently so would $|\xi|$. From (22) we see that N_ρ/N would approach 1 for any finite value of ρ. This means that the power transmitted through a finite area surrounding the wire would be zero if the total power were finite. Thus, the solution of Maxwell's equations loses its sense.

Though a single wire transmission line based on Sommerfeld's surface wave is favorable with regard to the losses effected by the conductor itself, the practical application is restricted to very high frequencies by reason of the large extension of the field. Below 3000 mc the size of low loss launching devices and the clearance area around the wire become too large. It will be shown in the next section how the extension of the field can be

reduced to such an amount that surface wave transmission lines become feasible also for lower frequencies.

III. WAVES ON A CONDUCTOR WITH MODIFIED SURFACE

The waves propagated in coaxial cables or in wave guides differ from Sommerfeld's surface wave in that the finite conductivity of the conductors is an undesired property rather than a necessary condition for the existence of the waves. The question arises: is it possible to modify the surface of a wire so that the conductivity is no longer important for the existence of a surface wave. It is evident that the field of a surface wave on such a guide would not extend as far in the radial direction as the field of Sommerfeld's wave; the latter wave is constrained to the wire only by reason of the finite conductivity while in the other case an additional constraint must be present which concentrates the field around the wire if the conductivity becomes infinite.

In answering the question we first consider the boundary conditions for a surface wave if there is no dissipation present.

An unattenuated surface wave of axial symmetry has field components which are also described by Eq. (1). The cylinder functions in question are the Hankel-functions $H_0^{(1)}$ and $H_1^{(1)}$ with positive imaginary argument. Hence

$$Z_0=H_0^{(1)};\quad Z_1=H_1^{(1)};\quad \gamma=j\gamma';\quad \gamma'^2=h^2-k^2. \tag{23}$$

The Hankel function $H_1^{(1)}$, for positive imaginary argument is negative real and the Hankel function $H_0^{(1)}$, for positive imaginary argument is negative imaginary. Hence E_z has a phase shift of 90° with respect to E_r and H_φ. The boundary conditions at the surface of a guide which propagates a wave of this type demand that E_z and H_φ and therefore their ratio

$$(E_z/H_\varphi)_{r=a'}=\frac{\omega\mu\gamma'}{k^2}\frac{H_0^{(1)}(j\gamma'a')}{H_1^{(1)}(j\gamma'a')}$$

$$\simeq -j(\mu/\epsilon)^{\frac{1}{2}}\frac{\gamma'^2a'}{k}\ln 0.89\gamma'a' \tag{24}$$

be continuous. a' is the radius at which the internal field and the field of the surface wave meet. Equation (24) states that at the surface of the guide the ratio $(E_z/H_\varphi)_{r=a'}$ must be positive imaginary.

Now we investigate whether the required ratio E_z/H_φ can be obtained by a suitable modification of the surface of a perfectly conducting wire. Consider a cylinder of radius $r=a'$ surrounding a wire which may have any periodic "micro" structure in the z-direction as indicated in Fig. 6. The space between the conductor surface and said cylinder may be partly or wholly filled with dielectric material. This space should be small compared with the wave-length, but large enough compared with the structure on the wire that at $r=a'$ the effect of this structure on the field is averaged. The region bounded

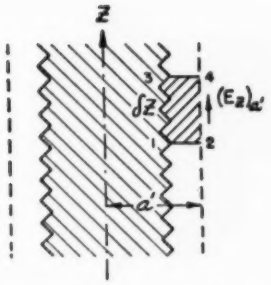


FIG. 6. Conductor with modified surface.

by the cylinder of radius a' is the "guide." With these assumptions the field in the guide can be considered quasi-stationary. The magnetic field is determined by the current I in the wire. The contribution of the electric z-component to the magnetic field can be neglected. The electric field consists in its major part of a potential field originated by the electric charges on the conductor and in its minor part of a curl-field induced by the magnetic field. Current I and charge Q per unit length are connected by the relation

$$\partial I/\partial z+\partial Q/\partial t=0. \tag{25}$$

Since the current distribution propagates in the z-direction with the same phase velocity as the surface wave, (25) can be written in the form

$$-jhI+j\omega Q=0 \quad \text{or} \quad Q=hI/\omega. \tag{26}$$

Faraday's induction law applied to the area 1–3–4–2 in an axial plane (see Fig. 6) gives:

$$\delta z\frac{\partial\Phi}{\partial t}=(E_z)_{a'}\delta z+\left(\int_1^2 E_r dr-\int_3^4 E_r dr\right). \tag{27}$$

The length δz, like the distance 1–2, must be small compared with the wave-length. Φ is the magnetic induction flux through a unit length of the area and is proportional to I. In the equation

$$\phi=LI, \tag{28}$$

L can be considered as an inductance per unit length. $(E_z)_{a'}$ in (27) is the longitudinal electric field component at $r=a'$. The integral of the radial electric field component between the points 1 and 2 is proportional to the charge density at 1. A radial induced component of the electric field can be neglected because the dimensions of the roughness of the conductor are assumed to be small compared with the dimensions of the area under consideration. Hence,

$$\int_1^2 E_r dr=Q/C \tag{29}$$

where C has the meaning of a capacitance per unit length. With (29) and (26) the difference between the two integrals in (27) can be written in the form

$$\int_1^2 E_r dr-\int_3^4 E_r dr=-\frac{1}{C}\frac{\partial Q}{\partial z}\delta z=j\frac{h^2}{\omega C}I\delta z. \tag{30}$$

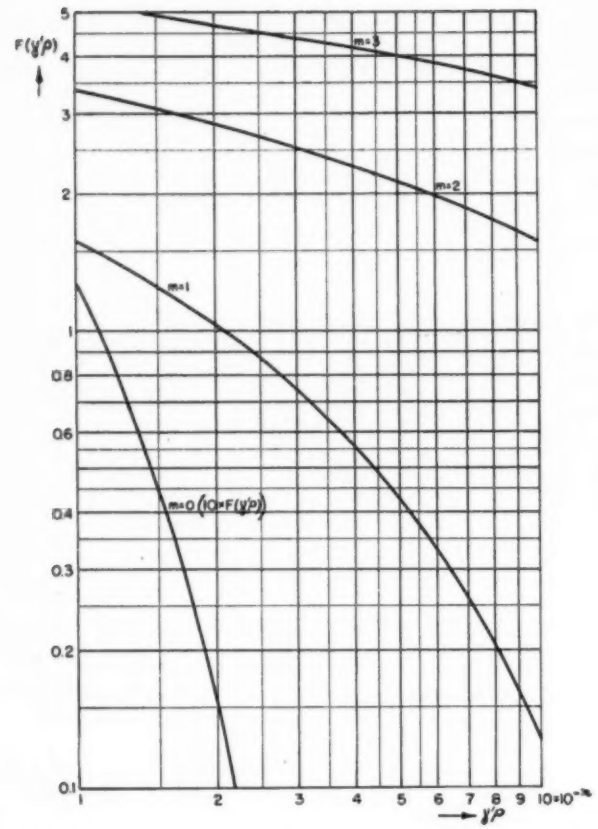


FIG. 7. Graph for determining the extension of the field on surface treated wires (see Eq. (37)).

With (28) and (30) inserted, Eq. (27) becomes:

$$j\omega LI\left(1-\frac{h^2}{\omega^2 LC}\right)=(E_z)_{a'}. \tag{31}$$

Assume for the moment that $(E_z)_{a'}$ is zero. Then, the conditions are like those in a coaxial line having the structure of our guide inclosed in a conductive tube of radius a'. The propagation constant h' of such a line is given by (31) for $(E_z)_{a'}=0$:

$$1-\frac{h'^2}{\omega^2 LC}=0 \quad \text{or} \quad h'^2=\omega^2 LC.^* \tag{32}$$

$\omega^2 LC$ in (31) can now be expressed by h', the propagation constant of a wave which would be translated if the surface wave guide were surrounded by a cylindrical conductor of radius a'. Thus Eq. (31) reads:

$$j\omega LI\left(1-\frac{h^2}{h'^2}\right)=(E_z)_{a'}. \tag{33}$$

As the magnetic field at $r=a'$ is $I/2\pi a'$, the ratio E_z/H_φ at $r=a'$ becomes:

$$(E_z/H_\varphi)_{r=a'}=j2\pi a'\omega L(1-h^2/h'^2). \tag{34}$$

* In truth L and C would change somewhat if the guide were surrounded by a conductor.

The surface wave requires according to (24) an E_z/H_φ having a positive imaginary value. Hence, h' must be greater than k. Because $h=(k^2+\gamma'^2)^{\frac{1}{2}}$ is greater than k, the condition for the existence of the wave considered is that *the guide, inclosed in a conducting cylinder, form a transmission line the phase velocity of which is smaller than the normal velocity in the surrounding dielectric.* Any surface modification of the wire which reduces the phase velocity is suitable for converting the wire into a surface wave guide. Threading of the wire or a dielectric layer on it, for example, accomplishes the required reduction in phase velocity. It may be mentioned that Sommerfeld's wave also is made possible by the fact that the phase velocity is reduced. In that case the reduction is effected by the finite conductivity of the wire. Several steps in the preceding development might require more detailed discussions for their justification. For the sake of brevity such discussions have been omitted.

After having established the existence of a non-radiating surface wave we investigate the extension of the field, still with the assumption that dissipation is negligible. As in the case of Sommerfeld's wave we determine first the amount of power which is transmitted outside a cylinder of radius ρ surrounding the wire. This power is given by (18) and since $H_1^{(1)}(j\gamma' r)$ is real,

$$N_\rho=2\pi AA^*(\epsilon/\mu)^{\frac{1}{2}}\frac{hk}{\gamma'^2}\int_\rho^\infty r[H_1^{(1)}(j\gamma' r)]^2 dr$$

$$=AA^*\pi(\epsilon/\mu)^{\frac{1}{2}}\frac{hk}{\gamma'^2}F(\gamma'\rho) \tag{35}$$

where

$$F(\gamma'\rho)=(\gamma'\rho)^2\left\{-\frac{2}{\gamma'\rho}jH_0^{(1)}(j\gamma'\rho)H_1^{(1)}(j\gamma'\rho)\right.$$

$$\left.-[H_0^{(1)}(j\gamma'\rho)]^2-[H_1^{(1)}(j\gamma'\rho)]^2\right\}. \tag{36a}$$

Hankel functions with imaginary argument are tabulated;[5] thus, the function $F(\gamma'\rho)$ can be calculated easily. In the range $\gamma'\rho<0.1$ the Hankel functions can be approximated by their zero developments with good accuracy. In this way the function $F(\gamma'\rho)$ becomes

$$F(\gamma'\rho)\simeq 8/\pi^2[-\ln 0.89\gamma'\rho-0.5] \quad \text{for} \quad \gamma'\rho<0.1. \tag{36b}$$

The function $F(\gamma'\rho)$ is plotted in Fig. 7. The parameter m indicates the decade in $\gamma'\rho$ to which each curve refers. In the case $m=0$ the function $10\times F(\gamma'\rho)$ is plotted in order to get the curve on the diagram.

The circle of radius ρ_p within which a certain percentage p of the total power of the surface wave travels is given by the equation:

$$p=1-\frac{F(\gamma'\rho_p)}{F(\gamma'a')}. \tag{37}$$

[5] E. Jahnke and F. Emde, *Tables of Functions*, (Dover Publications, New York, 1943).

Figure 8 shows the ratio ρ_p/a as a function of $\gamma' a'$ for $p=50$ percent, 75 percent, 90 percent and 99 percent. For a given diameter a' of the guide the field decreases faster with increasing γ'. Assume, for instance, $a'=1$ cm, $\gamma'=2\times10^{-2}$; then half the power is within a radius of about 6 cm, 75 percent of the power is within $\rho=16$ cm, 90 percent is within $\rho=33$ cm and 99 percent is within $\rho=86$ cm. γ' can be varied by modifying the conductor surface. Thus, we have the possibility of restricting the extension of the field within limits to any desired value. A part of the energy travels in the space between the conductor surface and a', the radius of the guide. But this part is so small within the considered γ' range, that it can be neglected. (See Section IV.)

IV. CONDUCTOR WITH DIELECTRIC COAT

The simplest method for reducing the phase velocity and the extension of the field is the application of a dielectric coat to a plain wire. This case will now be considered in detail. The field in the dielectric layer can be described by Bessel functions (J_0, J_1) and Neumann functions (N_0, N_1). In Eq. (1) Z_0 and Z_1 are to be replaced by

$$Z_0=J_0+bN_0; \quad Z_1=J_1+bN_1, \tag{38}$$

where b is determined from the boundary condition. γ and k are provided with the subscript i indicating that they refer to the insulating layer:

$$\gamma_i^2=k_i^2-h^2; \quad k_i=\omega(\epsilon_i\mu_i)^{\frac{1}{2}}. \tag{39}$$

Let a be the radius of the wire and a' the radius of the outer surface of the dielectric coat; then we obtain from the boundary condition $E_z=0$ at the wire of infinite conductivity:

$$b=-J_0(\gamma_i a)/N_0(\gamma_i a), \tag{40}$$

and the ratio E_z/H_φ at the surface of the coat, $r=a'$, becomes:

$$\left(\frac{E_z}{H_\varphi}\right)_{r=a'}=-j(\mu_i/\epsilon_i)^{\frac{1}{2}}\frac{\gamma_i}{k_i} \times\frac{J_0(\gamma_i a')N_0(\gamma_i a)-J_0(\gamma_i a)N_0(\gamma_i a')}{J_1(\gamma_i a')N_0(\gamma_i a)-J_0(\gamma_i a)N_1(\gamma_i a')}. \tag{41}$$

This equation can be simplified if we are interested only in conditions for which the phase velocity of the surface wave is slightly reduced. These conditions are established if either the dielectric layer is thin compared with the wire radius, or the layer is of the same order of magnitude as the wire radius, but the radius itself is very small compared with the wave-length. In the first case we can write

$$J_0(\gamma_i a)\simeq J_0(\gamma_i a')+\gamma_i J_1(\gamma_i a')(a'-a); \quad N_0(\gamma_i a')\simeq N_0(\gamma_i a')+\gamma_i N_1(\gamma_i a')(a'-a). \tag{42}$$

With these relations (41) becomes:

$$\left(\frac{E_z}{H_\varphi}\right)_{r=a'}\simeq j(\mu_i/\epsilon_i)^{\frac{1}{2}}\frac{\gamma_i^2}{k_i}(a'-a). \tag{43}$$

In the other case we use the zero representations of the cylinder functions. Thus (41) becomes:

$$\left(\frac{E_z}{H_\varphi}\right)_{r=a'}\simeq j(\mu_i/\epsilon_i)^{\frac{1}{2}}\frac{\gamma_i^2}{k_i}a'\ln\frac{a'}{a}. \tag{44}$$

Both Eqs. (43) and (44) are immediately obtained from Eq. (34). If the guide, consisting of the wire and the dielectric layer is thought to be surrounded by a metal tube of radius a', the inductance L per unit length and the propagation constant h' of the coaxial line formed by the guide and the tube are:

$$L=\frac{\mu_i}{2\pi}\ln\frac{a'}{a}; \quad h'=\omega(\epsilon_i\mu_i)^{\frac{1}{2}}=k_i. \tag{45}$$

With these quantities Eq. (34) leads to (44) which we consider more general since it yields, for thin layers, Eq. (43).

At $r=a'$ where the field in the guide and the field of the surface wave meet, $(E_z/H_\varphi)_{r=a'}$ must be continuous. Hence, by comparison of (24) and (44) we get the relation:

$$(\mu/\epsilon)^{\frac{1}{2}}\frac{\gamma'^2}{k}a'\ln 0.89\gamma' a'=-(\mu_i/\epsilon_i)^{\frac{1}{2}}\frac{\gamma_i^2}{k_i}a'\ln\frac{a'}{a} \tag{46}$$

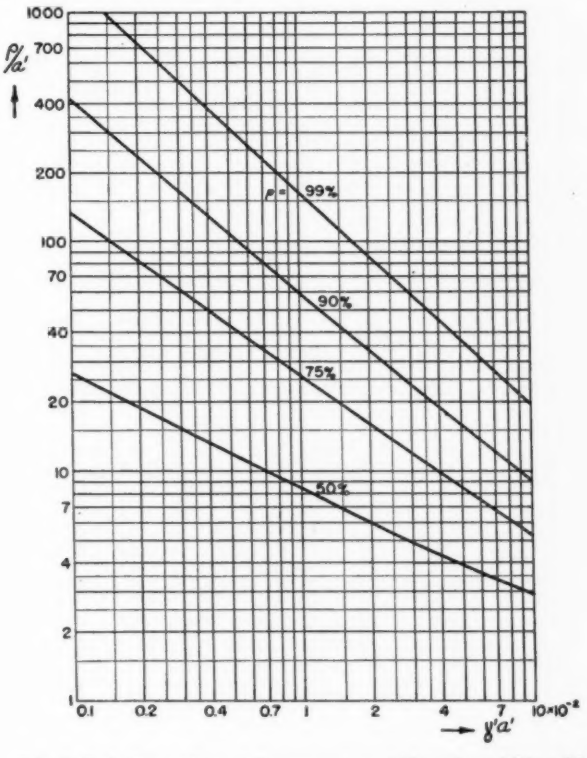


FIG. 8. Relative radius of the area around the wire within which 50 percent, 75 percent, 90 percent and 99 percent of the power is propagated in the case of a conductor with modified surface.

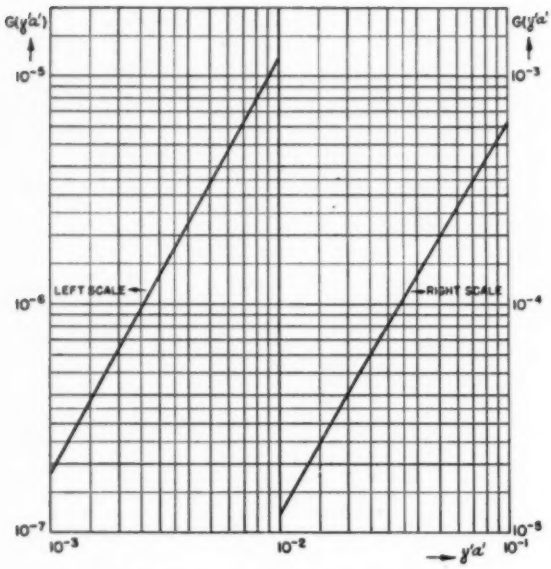


FIG. 9. Design curve in the case of dielectric coated wires (see Eq. (48)).

from which the ratio a'/a is determined:

$$\ln\frac{a'}{a}=-\frac{\epsilon_i}{\epsilon}\frac{\gamma'^2}{\gamma_i^2}\ln 0.89\gamma'a'. \tag{47}$$

As $\gamma_i^2=k_i^2-h^2=k_i^2-k^2-\gamma'^2$ (see Eqs. (39) and (23)) and $\gamma'^2\ll k^2$ Eq. (47) becomes for $\mu_i=\mu$:

$$\ln\frac{a'}{a}\simeq\frac{\epsilon_i}{\epsilon_i-\epsilon}\left(\frac{\lambda}{a'}\right)^2 G(\gamma'a')$$

with

$$G(\gamma'a')=-\left(\frac{\gamma'a'}{2\pi}\right)^2\ln 0.89\gamma'a'. \tag{48}$$

The function $G(\gamma'a')$ is plotted in Fig. 9.

We consider the following example: $a'=1$ cm; $\lambda=50$ cm. The radius within which 90 percent of the power of the surface wave is transmitted, $\rho_{90\%}=50$ cm. What is the required thickness of the dielectric layer if $\epsilon_i/\epsilon=4$?

From Fig. 8 we obtain the required value of $\gamma'a'$ for which $\rho_{90\%}=50$ cm. It is $\gamma'a'=1.21\times10^{-2}$. The corresponding value for $G(\gamma'a')$ is found with Fig. 9 to be 1.68×10^{-5}. Inserted in Eq. (48), these values give the thickness of the layer as 5.6×10^{-2} cm.

The next question to be answered is, how much power is transmitted inside the guide, i.e., within the dielectric coat. This power is determined by E_r and H_φ in the dielectric which are, in terms of the current I,

$$E_r=\frac{Q}{2\pi r\epsilon_i}=\frac{h}{2\pi\epsilon_i\omega r}I \text{ (see (26))}$$
$$H_\varphi=I/2\pi r. \tag{49}$$

Hence:

$$N_i=2\pi\int^{a'} rE_rH_\varphi^*dr=\frac{h}{2\pi\omega\epsilon_i}\ln\frac{a'}{a}II^*. \tag{50}$$

The power N of the surface wave itself is given by (35) with $\rho=a'$. For $F(\gamma'a')$ the approximation (36b) can be used; thus

$$N\simeq-\frac{8}{\pi}AA^*(\epsilon/\mu)^{\frac{1}{2}}\frac{hk}{\gamma'^4}(\ln 0.89\gamma'a'+0.5). \tag{51}$$

The amplitude A can be expressed in terms of the current. Since $(H_\varphi)_{r=a'}=I/2\pi a'$ we find from the field Eq. (1) with the substitutions of (23),

$$AA^*\simeq II^*\frac{\mu\gamma'^4}{16\epsilon k^2}. \tag{52}$$

With (50), (51), (52), (47) and (23) the ratio between N_i and N becomes:

$$\frac{N_i}{N}=-\frac{\epsilon}{\epsilon_i}\frac{\ln(a'/a)}{\ln 0.89\gamma'a'+0.5}$$
$$=\frac{h^2-k^2}{k^2(\epsilon_i/\epsilon)-h^2}\frac{\ln 0.89\gamma'a'}{\ln 0.89\gamma'a'+0.5}. \tag{53}$$

Within the range of $\gamma'a'$ between 10^{-3} and 10^{-1} the term $(\ln 0.89\gamma'a')/(\ln 0.89\gamma'a'+0.5)$ is smaller than 1.25. If we consider reductions in phase velocity of less than 3 percent and assume $\epsilon_i/\epsilon>2.4$, N_i/N is smaller than 6 percent. Thus N_i can be neglected in most practical cases.

V. ATTENUATION ON DIELECTRIC COATED CONDUCTORS

All preceding relations derived for a surface wave guide consisting of a conductor with modified surface are valid under ideal conditions where no dissipation is present. The effect of dissipation can be calculated in the usual manner with the assumption that the field distribution in an equiphase plane is approximately the same as in the dissipationless case.* This means that the losses are calculated for the undisturbed field. The power loss within the path from z to $z+dz$ is proportional to the total power

$$d\bar{N}=-2\beta\bar{N}dz, \tag{54}$$

where β is the attenuation factor in neper/cm.

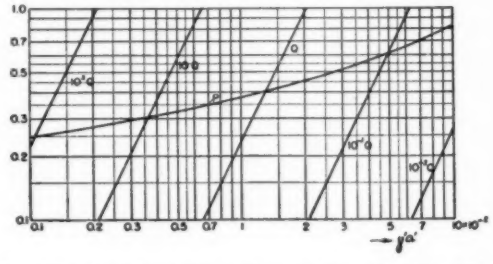


FIG. 10. Chart for determining P and Q (for Eq. (60)).

* This method could not be used in the case of Sommerfeld's wave for obvious reasons.

We consider only the case of a conductor with a dielectric layer. Then the loss consists of two parts, the conductivity loss and the dielectric loss. The conductivity loss, expressed in terms of the current I in the wire, is given by:

$$d\bar{N}_c = -\frac{1}{2\pi a}\left(\frac{\omega\mu_c}{2\sigma_c}\right)^{\frac{1}{2}} II^* dz \tag{55}$$

where μ_c and σ_c are the permeability and the conductivity of the wire. The dielectric loss is

$$d\bar{N}_i = -\frac{h^2}{2\pi\omega\epsilon_i}\ln\frac{a}{a}\tan\delta\, II^* dz$$

$$\simeq \frac{h^2\gamma'^2}{2\pi\omega\epsilon\gamma_i^2}\ln(0.89\gamma' a')\tan\delta\, II^* dz. \tag{56}$$

$\tan\delta$ is the loss factor of the dielectric material. The total power $\bar{N} = N + N_i$ is obtained from (51)–(53):

$$\bar{N} = -\frac{1}{2\pi}(\mu/\epsilon)^{\frac{1}{2}}\frac{h}{k}\left[\left(1+\frac{\gamma'^2}{\gamma_i^2}\right)\ln 0.89\gamma' a' + 0.5\right] II^*. \tag{57}$$

Equations (54)–(57) give the attenuation factor β_c caused by the conductivity losses and the attenuation factor β_i caused by the dielectric losses. Since $(\gamma'/\gamma_i)^2 \ll 1$, $h \simeq k$

$$\beta_c \simeq -\frac{1}{2a}\left(\frac{\omega\epsilon\mu_c}{2\sigma_c\mu}\right)^{\frac{1}{2}}\frac{1}{\ln 0.89\gamma' a' + 0.5}\ [\text{neper/cm}], \tag{58}$$

$$\beta_i \simeq \frac{1}{2}\frac{\epsilon}{\epsilon_i - \epsilon}\frac{\gamma'^2}{k}\left(1 - \frac{0.5}{\ln 0.89\gamma' a' + 0.5}\right) \times \tan\delta\ [\text{neper/cm}]. \tag{59}$$

With (58) and (59) the total transmission loss L in db per 100 feet becomes

$$L_{100\text{ ft.}} = \frac{P(\gamma' a')}{a(\lambda)^{\frac{1}{2}}} + \frac{\epsilon}{\epsilon_i - \epsilon}\tan\delta\frac{\lambda}{a'^2}Q(\gamma' a')[db] \tag{60}$$

with

$$P(\gamma' a') = -1.33\times 10^4(\epsilon/\mu)^{\frac{1}{2}}\left(\frac{\pi\mu_c}{\sigma_c\mu}\right)^{\frac{1}{2}}\frac{1}{\ln\gamma' a' + 0.38},$$

$$Q(\gamma' a') = 2.11\times 10^3\left(1 - \frac{0.5}{\ln\gamma' a' + 0.38}\right)(\gamma' a')^2. \tag{61}$$

For a coated copper wire stretched in air,

$$P(\gamma' a') = 1.60\frac{1}{-\ln\gamma' a' - 0.38}. \tag{62}$$

The functions $P(\gamma' a')$ (for copper) and $Q(\gamma' a')$ are plotted in Fig. 10.

FIG. 11. Sketch of a surface wave transmission line.

For the example considered in the preceeding section ($\lambda = 50$ cm, $\gamma' a' = 1.21\times 10^{-1}$, $a = 0.94$ cm, $a' = 1.00$ cm, $\epsilon_i/\epsilon = 4$) we obtain from Fig. 10: $P = 0.38$, $Q = 0.29$. The loss due to the conductivity is given by the first term in (60) and is 0.057 db/100 ft. Assuming $\tan\delta = 10^{-3}$, the dielectric loss, given by the second term of (60), is 0.005 db/100 ft. Thus the total loss is 0.062 db/100 ft. As may be seen from Fig. 4 for a plain wire of the same radius guiding a Sommerfeld surface wave, the loss would be only half as large. However, the extension of the field would be so high that the conditions for undisturbed propagation and efficient excitation could not be achieved. Figure 5 shows that for $a = 1$ cm and $\lambda = 50$ cm, $\rho_{50\%} = 34$ cm and $\rho_{75\%} = 200$ cm. The corresponding radii for the coated wire given in Fig. 8 ($\gamma' a' = 1.21\times 10^{-1}$) are $\rho_{50\%} = 7.6$ and $\rho_{75\%} = 22.5$ cm; this means they amount to only 22 percent and 11 percent of the other radii.

VI. LAUNCHING OF SURFACE WAVES

In order to excite the desired wave mode with high efficiency a launching device is required which develops within a cross section a field having a structure similar to that of the surface wave. The structure of the fields of both types of surface wave considered in this paper is, within a large area around the conductor, somewhat like in a coaxial cable. The radial electric field component and the circular magnetic field component decrease with $1/r$ as can be seen from the zero representation of the Hankel function $H_1^{(1)}$. Furthermore the ratio E_r/H_φ is very close to $(\mu/\epsilon)^{\frac{1}{2}}$. As the major part of the energy of the waves travels within the area in which the $1/r$ decrease of the field strength is valid, the waves can be launched with good efficiency by means of the device sketched in Fig. 11. The outer conductor of the coaxial feed line is gradually expanded in the form of a horn and the inner conductor of the line is connected to the surface wave guide. The device can be considered as a tapered coaxial line. The arrangement on the receiving end is the same. In the case of Sommerfeld's surface wave the area in which the field should be simulated is very large especially for lower frequencies. Hence, the efficient excitation of this wave would require an extremely large horn. In the other case, where the extension of the field is reduced by means of a suitable modification of the conductor surface, the dimensions of launching devices can be kept within reasonable limits.

The efficiency of the launching device of Fig. 11 can be determined approximately if the opening of the horn does not too much exceed the area within which the field decreases with $1/r$. In this case, considering the receiving

end, that portion of the wave will be received which falls within the area of the horn. If the radius of the horn is equal to $\rho_{75\%}$, for instance, about 75 percent of the energy will be received. Because of the reciprocity theorem the efficiency of the horns for transmission and reception must be the same.

The tapering of the horns is important too, insofar as in the case of large angles higher modes may develop inside the horns, and besides, the wave front at the end of the horns may deviate too much from the required plane wave front. It is desirable to have the surface modification started within the horns, in order to have the surface wave build up therein.

In the case of a surface wave guide consisting of a wire with modified surface, the efficiency of a horn can be determined with the aid of Fig. 7. If the parameter γ' and the radius a' of the guide are known $F(\gamma'a')$ can be obtained. Let b be the radius of the horn at the open end; then the efficiency is given by $1-F(\gamma'b)/F(\gamma'a')$ where $F(\gamma'b)$ is also obtained from Fig. 7.

VII. EXAMPLES

The set-up considered here is shown in Fig. 11. The surface wave guide consists of an enameled wire of 120-foot length having a diameter of 0.2 cm ($a=0.1$) and an enamel coat of 0.005 cm thickness. The dielectric constant and the power factor of the enamel are $\epsilon_i/\epsilon=3$ and $\tan\delta=8\times10^{-3}$, for low frequencies. $\tan\delta$ will probably be much larger for the high frequencies under consideration; however, having no better data available, we will use the quoted figure. The diameter of the horns is $2b=33$ cm.

First we determine the loss of the wire, assuming a frequency of 3300 mc. With (48) we find, $G(\gamma'a')=4.04\times10^{-6}$. From Fig. 9 we obtain the corresponding $\gamma'a'=5.6\times10^{-3}$; Fig. 10 gives for $P(\gamma'a')$ and $Q(\gamma'a')$ the values $P=0.33$, $Q=0.072$. These values inserted in (60) deliver the transmission loss per 100 feet as 1.35 db. Hence the loss for 120 feet is 1.62 db.

In order to determine the loss due to the launching device we determine $F(\gamma'a')$ and $F(\gamma'b)$ from Fig. 7, $F(\gamma'a')=4.9$, $F(\gamma'b)=0.17$. The efficiency of each horn is then 95.6 or the loss is 0.20 db. As both horns together contribute 0.40 db, the total theoretical loss of the transmission line becomes 2.0 db. The measured loss was 2.3 db. It should be mentioned that the loss of the setup does not vary more than 0.1 db within the frequency range from 1500 to 3400 mc.

As another example we consider a transmission line with the same launching device and an enameled wire of 600-foot length and 0.322 cm. diameter. The thickness of the insulation is 0.0025 cm. We calculate the loss for a frequency of 1600 mc. From (48) we obtain $G(\gamma'a')=0.75\times10^{-7}$ and from Fig. 9, $\gamma'a'=2.4\times10^{-3}$. The corresponding values for $P(\gamma'a')$ and $Q(\gamma'a')$ are found in Fig. 10. They are $P(\gamma'a')=0.28_3$ and $Q(\gamma'a')=0.013_1$. If these values are inserted in (60) the loss for the length of 600 feet is 2.65 db. The loss contributed by one horn is obtained with Fig. 7 which gives the values $F(\gamma'a')=4.6$ and $F(\gamma'b)=0.88$. Hence the efficiency is 80.9 percent or the loss 0.9_3 db. As the two horns contribute 1.8_6 db, the total loss of the transmission line becomes 4.5 db. The measured loss was approximately 5 db. It may be mentioned that the wire was supported at 100 feet intervals by strings about 6 feet above the ground.

VIII. ACKNOWLEDGMENT

Thanks are due Mr. A. Meyerhoff for his invaluable assistance in the preparation of the manuscript and Mr. C. Sharp and Mr. L. Battersby for valuable suggestions and for their cooperation in the performance of the experiments of which part of the results are quoted in this paper.

IX. APPENDIX

Derivation of Eq. (12):

If we write for ξ and η

$$\xi=|\xi|e^{j\alpha}=|\xi|e^{-j(\pi/4+x)};\quad \eta=|\eta|e^{j3\pi/4}\ (-\pi<x<\pi) \tag{A1}$$

(9) becomes

$$|\xi|e^{-jx}\left[\ln|\xi|-j\left(\frac{\pi}{4}+x\right)\right]=-|\eta|. \tag{A2}$$

This equation can be split into two equations if the real and imaginary parts are separated:

$$|\xi|\ln|\xi|\cos x\cdot\left(1-\frac{\pi/4+x}{\ln|\xi|}\tan x\right)=-|\eta|; \tag{A3}$$

$$\tan x=-\frac{\pi/4+x}{\ln|\xi|}. \tag{A4}$$

Inserting (A4) into (A3),

$$|\xi|\ln|\xi|\cos x\cdot(1+\tan^2 x)=-|\eta| \quad \text{or} \quad \frac{|\xi|\ln|\xi|}{\cos x}=-|\eta|. \tag{A5}$$

Within the range considered, $|\eta|<10^{-3}$, and $|\xi|\ll1$ (because $\gamma\ll k$). As $|\cos x|$ is certainly ≤1, $|\xi|$ will be smaller than 1.1×10^{-4} and $-\ln|\xi|>9$. From (A4) we see that for $-\ln|\xi|>9$, x becomes small:

$$x\simeq\frac{\pi}{4}\frac{1}{-\ln|\xi|-1}\left(<\frac{\pi}{32}\right). \tag{A6}$$

For such small angles $\cos x$ in (A5) can be replaced by 1. Thus the Eqs. (12) are obtained.

A Note on Control Area

THOMAS M. STOUT
Electrical Engineering Department, University of Washington, Seattle, Washington
(Received May 29, 1950)

Control area, or the integrated error for a step function input, has been used as a measure of control system performance. For systems described by

$$\frac{\theta_0(s)}{\theta_i(s)}=e^{-sT_d}\frac{(\tau_1 s+1)(\tau_2 s+1)\cdots(\tau_m s+1)}{(T_1 s+1)(T_2 s+1)\cdots(T_n s+1)},$$

it is proved that the control area is

$$A=T_d+\sum_{i=1}^{n} T_i-\sum_{i=1}^{m}\tau_i.$$

A simple method of evaluation is presented and some limitations of the control area criterion are mentioned.

A COMMON measure of the response of a control system[1–3] to a unit step function is the "control area," shown as the shaded area in Fig. 1. The control area is defined mathematically as the integral of the error with respect to time from zero to infinity, the error being defined as the difference between input and output. An optimum adjustment of the system parameters is obtained when the control area is a minimum.

Evaluation of the control area in terms of the system parameters by conventional methods is a difficult task if the system is complicated. However, for systems described by the equation

$$\frac{\theta_0(s)}{\theta_i(s)}=\frac{(\tau_1 s+1)(\tau_2 s+1)\cdots(\tau_m s+1)}{(T_1 s+1)(T_2 s+1)\cdots(T_n s+1)} \tag{1}$$

$$=\frac{P(s)}{Q(s)}, \tag{2}$$

subject to a step function of magnitude M, it will be shown that the control area is simply

$$A=1/M\int_0^\infty e(t)dt$$

$$=(T_1+T_2+\cdots+T_n)-(\tau_1+\tau_2+\cdots+\tau_m). \tag{3}$$

In these equations, θ_0, θ_i, and e are standard symbols[4] for the output, input, and error, respectively, functions of s are the Laplace transforms of the corresponding functions of time, and $\tau_1\cdots\tau_m$, $T_1\cdots T_n$ are time constants of the system. Equation (1) describes systems with any number of "lag" or "lead" (delay or anticipatory) elements having zero steady-state error for a step function input. If the steady-state error is not zero, the control area as defined above is infinite and therefore useless as a measure of system performance.

PROOF

By means of Eq. (2), the error may be expressed in terms of the input by the equation

$$e(s)=\left[\frac{Q(s)-P(s)}{Q(s)}\right]\theta_i(s). \tag{4}$$

If the input is a unit step function whose Laplace transform is $1/s$, the transform of the error is

$$e(s)=\frac{Q(s)-P(s)}{sQ(s)}. \tag{5}$$

Denote by $I(t)$ the indefinite integral of the error

$$I(t)=\int e(t)dt. \tag{6}$$

Since $I(0)$ is necessarily zero, the Laplace transform of $I(t)$ is

$$I(s)=e(s)/s. \tag{7}$$

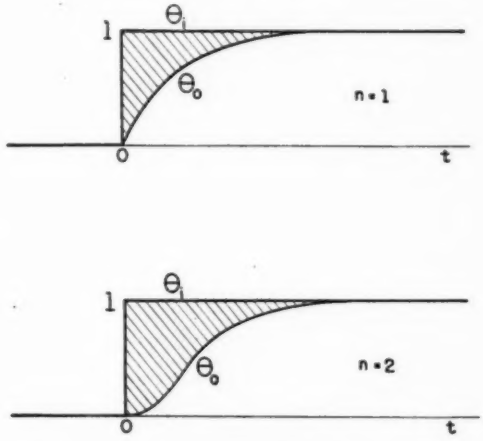


FIG. 1. Response and control area for simple control systems.

[1] R. C. Oldenbourg and H. Sartorius, *The Dynamics of Automatic Controls* (A.S.M.E., New York, 1948), pp. 65–67.
[2] C. Mack, Phil. Mag. **40**, 922–928 (September, 1949).
[3] D. P. Eckman, *Principles of Industrial Process Control* (John Wiley and Sons, Inc., New York, 1945), p. 161.
[4] G. S. Brown and D. P. Campbell, *Principles of Servomechanisms* (John Wiley and Sons, New York, 1948).

The control area may now be found from Eq. (7) by means of the final value theorem,[5,6] giving

$$A=\lim_{t\to\infty}[I(t)]$$

$$=\lim_{s\to 0}[sI(s)]$$

$$=\lim_{s\to 0}[e(s)]. \tag{8}$$

Substituting the expression for $e(s)$ from Eq. (5) into Eq. (8) gives

$$A=\lim_{s\to 0}\left[\frac{Q(s)-P(s)}{sQ(s)}\right], \tag{9}$$

which is indeterminate, since $Q(0)=P(0)=1$. Differentiating numerator and denominator, we obtain

$$A=\lim_{s\to 0}\left[\frac{Q'(s)-P'(s)}{sQ'(s)+Q(s)}\right] \tag{10}$$

$$=\lim_{s\to 0}[Q'(s)-P'(s)]. \tag{11}$$

At this point it will be convenient to write

$$P(s)=\prod_{i=1}^{m}p_i$$

$$Q(s)=\prod_{i=1}^{n}q_i, \tag{12}$$

where $p_i=\tau_i s+1$ and $q_i=T_i s+1$. The derivative of $P(s)$ or $Q(s)$ with respect to s is the sum of all the products formed by multiplying the derivative of one factor by all the other factors, or

$$P'(s)=\sum_{j=1}^{m}\tau_j\Big[\prod_{i=1}^{j-1}p_i\Big]\Big[\prod_{j+1}^{m}p_i\Big]$$

$$Q'(s)=\sum_{j=1}^{n}T_j\Big[\prod_{i=1}^{j-1}q_i\Big]\Big[\prod_{j=1}^{m}q_i\Big]. \tag{13}$$

Since p_i and q_i approach one as s approaches zero, we get

$$\lim_{s\to 0}P'(s)=\tau_1+\tau_2+\cdots+\tau_m$$

$$\lim_{s\to 0}Q'(s)=T_1+T_2+\cdots+T_n \tag{14}$$

and the proof of Eq. (3) is complete.

In systems characterized by "dead time," the output response does not begin immediately but starts at a later time, $t=T_d$. In this case,

$$\frac{\theta_0(s)}{\theta_i(s)}=\frac{e^{-sT_d}P(s)}{Q(s)} \tag{15}$$

and a similar proof will show that now

$$A=T_d+\sum_{i=1}^{n}T_i-\sum_{i=1}^{m}\tau_i. \tag{16}$$

This conclusion is also evident from the previous result upon consideration of the nature of the response.

DISCUSSION

Using conventional methods, the error would be written

$$e(t)=\sum_{i=1}^{m}K_i e^{-t/T_i}, \tag{17}$$

where the coefficients K_i are functions of $\tau_1\cdots\tau_m$ and $T_1\cdots T_n$, and the control area found by direct integration. The coefficients may be determined from the initial conditions* or from Eq. (5) by the use of partial fractions or residues. For the simple first- and second-order systems, whose responses are shown in Fig. 1, the error is

$$e_1(t)=e^{-t/T_1} \tag{18}$$

$$e_2(t)=\frac{T_1}{T_1-T_2}e^{-t/T_1}+\frac{T_2}{T_2-T_1}e^{-t/T_2}. \tag{19}$$

In these simple cases, the validity of Eq. (3) is easily demonstrated. However, the algebraic difficulties encountered in treating the general case by this method are staggering; the direct attack made possible by the Laplace transformation is simpler and more successful.

If $P(s)$ and $Q(s)$ should be given in the form

$$\begin{aligned}P(s)&=B_m s^m+B_{m-1}s^{m-1}+\cdots+B_1 s+1\\ Q(s)&=A_n s^n+A_{n-1}s^{n-1}+\cdots+A_1 s+1\end{aligned}, \tag{20}$$

the control area can be found directly without solving for the individual time constants. Expansion of the factored forms of $P(s)$ and $Q(s)$ used in Eq. (1) will show that

$$\begin{aligned}B_1&=\tau_1+\tau_2+\cdots+\tau_m\\ A_1&=T_1+T_2+\cdots+T_n.\end{aligned} \tag{21}$$

The control area is therefore given by A_1-B_1 directly. Using the control area criterion, the objective of system design would be to minimize (A_1-B_1). It should be remembered that allowable values of A_1 will be fixed by stability considerations, as given by Routh's stability criterion.[6,7] In particular, if $A_1\leqq 0$, the system will be unstable.

[5] R. V. Churchill, *Modern Operational Mathematics in Engineering* (McGraw-Hill Book Company, Inc., New York, 1944).

[6] M. F. Gardner and J. L. Barnes, *Transients in Linear Systems* (John Wiley and Sons, Inc., New York, 1942).

* For a unit step function, the error is equal to one and the first $n-1$ derivatives with respect to t are equal to zero.

[7] E. J. Routh, *Advanced Rigid Dynamics* (Macmillan Company, Ltd., London, 1930).

If the time constants are complex, Eq. (3) is still applicable. In physical systems complex time constants occur in conjugate pairs, such as

$$T=1/(a+j\omega_n),$$
$$\bar{T}=1/(a-j\omega_n),$$

where a is a damping factor and ω_n is the natural frequency. For such a pair of complex time constants,

$$T+\bar{T}=2a/(a^2+\omega_n^2). \tag{22}$$

This result indicates nicely one limitation of control area as a criterion for system design, since the minimum value of $T+\bar{T}$ is obtained when the damping is zero. Other systems, with equally undesirable responses, can be imagined for which the control area might be zero or even negative. These observations indicate the need for simultaneous consideration of the relative stability of the system. In systems for which the control area is a suitable measure of response, the methods of evaluation proposed in this note may be useful.

A Small Electron Microscope

JOHN H. REISNER AND EDMUND G. DORNFELD
Radio Corporation of America, Camden, New Jersey
(Received May 25, 1950)

A new electron microscope of greatly simplified design has been constructed. The instrument has a resolution of 100 Angstroms, yet is small enough to be mounted on a desk or laboratory table. Magnetic lenses are energized by permanent magnets. An accelerating voltage of 50 kilovolts is varied over a narrow range to provide means of focus. Images may be photographed on 2×2 plates. Plates may be changed in a minute, and specimens in fifteen seconds. Direct magnifications of 1500, 3000 and 6000 are possible depending on pole pieces used. Photographic images may be enlarged up to ten times. The instrument is free of x-radiation and completely interlocked and provided with shorting devices for protection from high voltage.

INTRODUCTION

ALL electron microscopes have been characterized heretofore by large size and weight, and usually by complexity of design. These characteristics are a result of extremely great operational flexibility of the instruments. In the early days of electron microscopy, such flexibility was essential to the determination of the best means of utilizing the optical components and to the development of new techniques. Complexity is still a necessary evil of the highest resolution instruments, and of those engaged in an extremely wide range of problems. However, many of the uses to which the electron microscope is now put[1] require only the moderately high resolution which may be realized on comparatively simple instruments.

A first step in the direction of smaller, simpler microscopes was made several years ago,[2] culminating in a well-known commercial instrument.[3] Additional simplification is possible by utilizing permanent magnets to replace coil windings, cables, control circuits, and power supplies. Unfortunately, the use of permanent magnets fixes the magnification for a given set of lens pole pieces. However, at the same time, alignment problems largely disappear so that transverse adjustments may be omitted, along with the attendant vacuum difficulties. By being content with ultimate resolutions of 100 Angstroms or less, the condenser lens may be omitted and its beam control function delegated to an aperture. By limiting direct magnification to 3000 or 6000 times, the column may be shortened, the volume of the vacuum system greatly reduced, and in turn, the vacuum diffusion pump may be reduced in size. Utilization of new materials such as silicone diffusion pump oil, combined with a new treatment of old valving and stage problems, have resulted in the greatly simplified electron microscope shown in Figs. 1 and 2.

DIMENSIONAL CONSIDERATIONS

The first step toward simplification is the reduction of size. The minimum dimensions for the microscope optical system are determined primarily by the highest magnification required, and by the focal lengths of the lenses. For the highest magnification, the lenses are so powerful that the magnification may be expressed to good approximation by the expression $M\simeq p/f$, where "M" is the magnification, and "p" and "f" are the image and focal distances of a lens. It is simply demonstrated that the most favorable utilization of the column length "L" is that where both lenses have the same image distances, i.e., where $p_1=p_2\cong L/2$, since f is very small compared to p. The minimum required column length then becomes $L_{\min}\cong 2(M_t f_1 f_2)^{\frac{1}{2}}$, where "$M_t$" is the total magnification, and "f_1" and "f_2" are the focal lengths of the objective and projector lenses respectively. For the microscope described, $f_1=0.25$ cm, and $f_2=0.20$ cm, when the accelerating voltage is 50 kilovolts and the over-all magnification is 6000 times. Under these circumstances, the minimum length turns out to be about 35 centimeters or 14 inches. It is inter-

[1] R. W. G. Wyckoff, *Electron Microscopy Technique and Aplications* (Interscience Publishers, Inc., New York, 1949), pp. 22–25, Chapters VI and VII.
[2] V. K. Zworykin and J. Hillier, J. App. Phys. **14**, 658 (1943).
[3] P. C. Smith and R. G. Picard, Radio News **32**, 41 (1944).

esting that a saving of 10 centimeters in length could be accomplished by limiting the top direct magnification to 3000 times.

The distance from the specimen to the illuminating source is also limited by magnification. The higher the magnification, the smaller must become the distance in order to maintain illumination. Unfortunately, this distance also influences resolution adversely as it is decreased. The actual values of the distances between optical parts are shown in Fig. 3. It is seen that the distance between the objective and projector lens gaps is substantially smaller than $L/2$. The distance is actually $0.37L$, decreasing the maximum magnification possible in length L by six percent, which is negligible. Design considerations governing the lens energizing system and the mechanical construction of the pole pieces determine the spacing between the two lens gaps.

The short space below the column provided for the opening of the plate holder door, and the space above the column necessary to provide protection against high voltage flash-over, raise the total length of the microscope assembly to 30 inches from base to top of the hood. The diameter of the body of the microscope is determined by the lens. Smaller diameters than 4.75 inches result in the operation of the permanent magnets at too high a demagnetization coefficient.

THE OPTICAL SYSTEM

The optical elements of the microscope are all fixed, with the exception of the electron gun assembly which is externally adjustable. This need for adjustment in the gun system is occasioned by the demountable nature of the filament which must be replaced from time to time. Since filaments differ somewhat in shape and in the position of the point of maximum electron emission, adjustment must be provided for the gun assembly. The gun assembly comprises a filament-grid assembly, rigidly attached to an anode. The anode is positioned with respect to the other optical elements by means of two screws loaded by a return spring, Figs. 4 and 5. The filament-grid portion of the gun is physically described[4] and its operation is discussed[5] in the literature. The electrodes are connected in a self-biased circuit, giving rise to a very intense solid cone of illumination. It is the function of the transverse gun adjustment controls to select that element of the cone of illumination which is parallel to the optical axis of the system and position it on that axis.

The alignment reference of the entire optical system is the bore of the lens energizing spool. The projector pole piece assembly and the lower element of the objective pole piece are positioned in the spool hole to the tolerance of the machining, nominally 0.0003 inch. The position of the upper element of the objective lens must be experimentally determined and fixed when the instrument is initially put in operation. Early embodiments of the instrument positioned the upper lens ele-

FIG. 1. The small electron microscope column completely assembled.

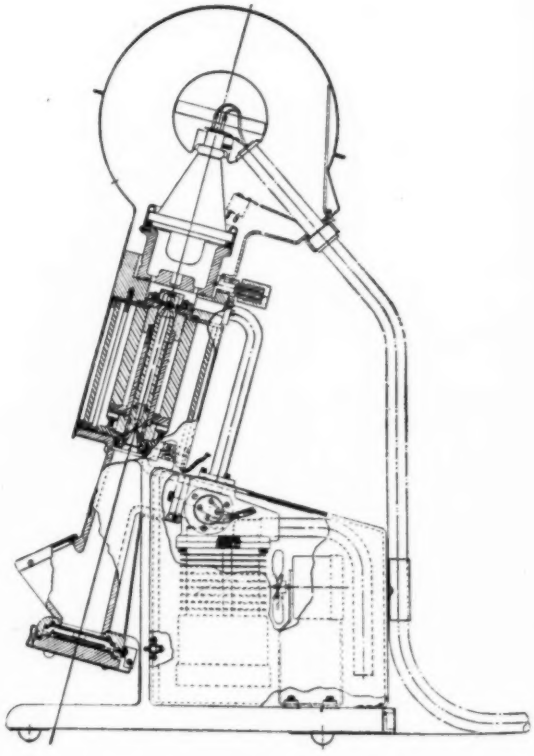


FIG. 2. A cross-sectional view of the electron microscope column employing permanent magnets.

[4] J. H. Reisner and R. G. Picard, Rev. Sci. Inst. **19**, 556 (1948).
[5] J. Hillier and S. G. Ellis, J. App. Phys. **20**, 700 (1949).

ment by its accurate fit into the energizing spool bore. However, the lateral positioning of the objective pole piece is extremely critical. The nominal build-up of fabrication tolerances amounting to 0.0003 inch can provide an image shift at the final screen of 0.15 inch per kilovolt change in accelerating voltage at fifty kilovolts and at a magnification of 6000 times. So large a shift is objectionable. To permit accurate electrical alignment of the upper objective element, a reference yoke is provided to assure position of that element. The yoke is transversely adjustable by means of four tapered screws (Fig. 5), which also serve to lock it in position so that the lens element may be removed and replaced without affecting alignment. Once aligned, the adjustment for the yoke is correct for all sizes and combinations of objective elements.

The objective aperture is referenced to the hole of the upper objective lens element, and is positioned 0.058 inch below the upper surface of the pole piece. The specimen is positioned 0.023 inch above the upper surface of the pole piece.

An aperture placed 0.434 inch above the specimen is used to exercise the beam control function normally accomplished by the condenser lens system of larger microscopes. Actually, the lack of a condenser lens ultimately limits the resolution of the instrument to that commensurate with the angular aperture of the electron gun. The disk-shaped aperture fits accurately in a hole of a heavy aperture holder, which is in turn referenced to the same yoke as that which aligns the upper lens element of the objective (Fig. 5). The massive aperture holder firmly held in the yoke does not move under the heating from the gun bombardment and shields the specimen rod from the heating and contamination of the gun.

Since permanent magnets energize the magnetic lenses, change in focal length may be accomplished only by changing the accelerating voltage applied to the microscope gun. The range of voltage necessary to provide focus for a wide range of specimens and holders is of the order of 500 volts at 50 kilovolts. A wider range is desirable for checking alignment and for checking the gap spacing of the objective pole pieces. A four-kilovolt swing produced through continuous coarse and a fine control has been found most suitable.

In order to assure the occurrence of a focused image at given voltage, e.g., 49 kilovolts, some variation in one of the focal lengths or object distances must be provided by mechanical means to compensate for differences in the build-up of mechanical tolerances of components and variations in the strength of energizing magnets. This required adjustment is provided in the gap spacing of the objective lens. The upper pole face of the objective pole piece seats against the energizing spool, and is, therefore, fixed with respect to the specimen. The lower pole of the objective, and the complete projector pole piece assembly are rigidly attached to one another. Motion of the lens group along the axis

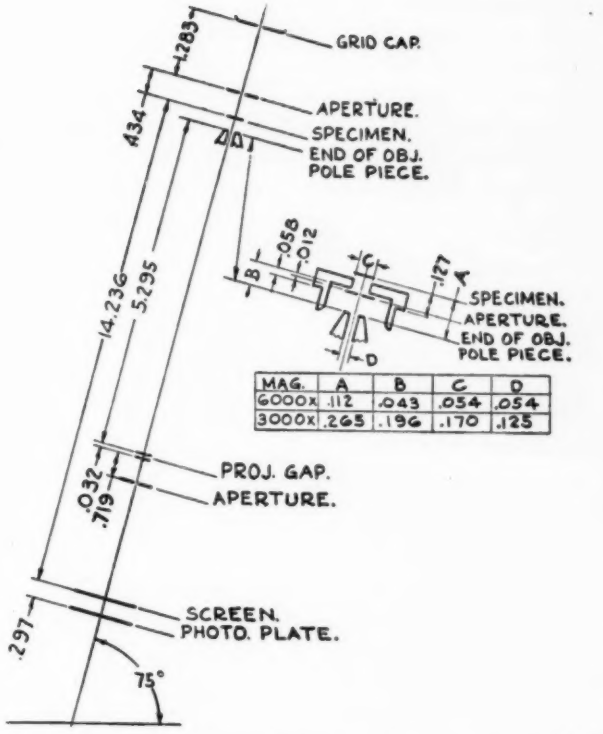


FIG. 3. A dimensional diagram of the optical system showing lens dimensions for magnifications of 3000× and 6000×.

of the lens energizing spool is possible, and permits adjustment of the spacing of the objective pole piece gap. The axial movement is possible in fine increments with reproducibility by means of the screw thread which advances the projector pole piece as it is turned.

Changing the gap separation shifts the principal planes of the lens in the direction of the change. Since focal length is relatively independent of gap separation, the change in focal length due to gap increase is small compared to shift of the principal plane.[6] The result is that the lens system may be positioned so as to bring the focal plane of the objective into coincidence with the specimen plane for any of the available range of voltages by simply changing the objective gap.

In the initial adjustment of the instrument, too small a gap will result in a focal plane on the gun side of the specimen which manifests itself in an approach toward focus at the low voltage end of the focus controls. Such evidence indicates that the pole piece gap must be increased by turning the projector pole piece counterclockwise. Likewise, approaching focus at the high voltage end of the focus control means that the objective gap is excessive, which must be corrected by a clockwise rotation of the projector pole piece. The amount of the rotation depends upon the objective pole piece in use and upon the accelerating voltage. The required pole piece rotation may be estimated

[6] Zworykin, Morton, Ramberg, Hillier and Vance, *Electron Optics and The Electron Microscope* (John Wiley and Sons, Inc., New York, 1945), p. 153.

from the relation:

$$\frac{1}{f}=\frac{K}{V}$$

taken from the magnetic lens formula, where "f" is the focal length, "V" the accelerating voltage, and "K" a constant depending upon factors fixed for this application. From this relation, it is apparent that:

$$\frac{df}{f}=\frac{dV}{V}.$$

In powerful lenses, used at large magnification as herein, it may be assumed that $f\cong q$, the object distance, so that for small increments:

$$\frac{\Delta q}{\Delta V}=\frac{q}{V}.$$

For the 3000× objective of Fig. 6, $q=0.138$ inch, and $V=50{,}000$ volts.

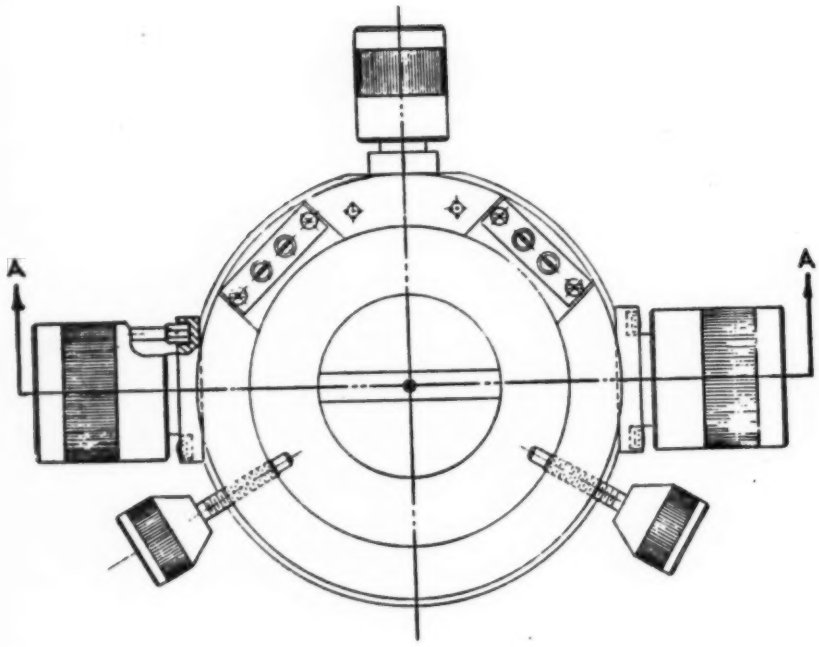
$$\frac{\Delta q}{\Delta V}\cong 0.0028 \text{ mils/volt.}$$

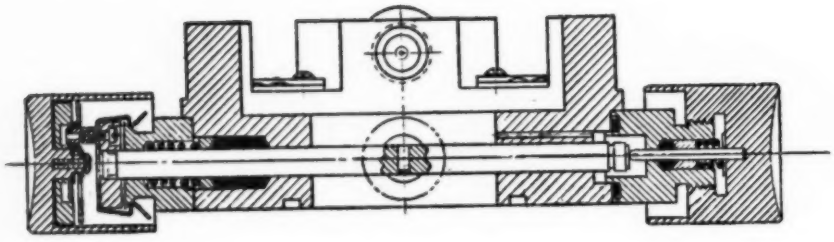
Since the gap change results from turning a 32 thread,

$$\frac{\Delta \theta}{\Delta V}\cong 0.033 \text{ degree/volt,}$$

where "θ" is the rotation of the pole piece required to advance it. Hence, if the voltage at which focus occurs is to be increased from 46 to 49 kilovolts, the pole piece must be turned 99 degrees counterclockwise.

The necessity for cleaning off contamination deposited on pole pieces during operation is the reason for the demountable construction of lens elements. Such con-





SECTION A-A

FIG. 4. Cross-sectional view of specimen and electron gun controls.

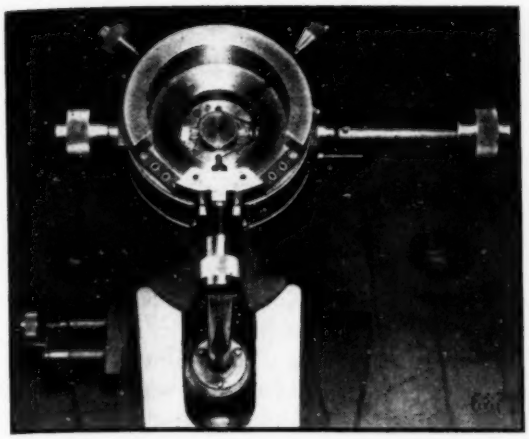


FIG. 5. Rear view of the upper end of the column with the gun assembly removed to show philister head screws used to align the upper element of the objective lens. The large circular body in the center of the chamber is the holder for the "condenser" aperture. The two knobs at the top of the picture adjust the position of the gun assembly.

tamination occurs in all objectives, but negligibly in projector lenses where the beam intensity is very low and the pole faces are completely shielded from the gun. The objective lens of the described microscope is easily demounted for cleaning and replacing apertures. The upper element may be removed through the specimen chamber, at which time the lower element of the objective pole piece is exposed for inspection and minor cleaning. Occasionally, it is necessary to remove the lower objective element. To do this, the projector pole piece is screwed out of its adjustment threads and the assembly pushed down and out of the lens energizing spool. The plate door and shutter should be open to admit the fingers to guide the pole piece out of the viewing chamber. Since the projector pole piece does not contaminate to any appreciable extent, its two pole piece elements are forced permanently onto its nonmagnetic spacer. Large holes in the spacer wall permit access to the pole faces for blowing out dust and for permitting minor cleaning.

The lower objective pole piece element is held tightly to the projector pole piece by set screws which pull down on a conical surface in such a way that the two iron pieces register along an inside surface. Such close registry minimizes stray fringe flux which might influence the electron beam.

Apertures are used (1) to restrict the beam striking the specimen, (2) to intercept electrons scattered from the specimen and thus increase image contrast, and (3) to prevent unused electrons from the powerful projector lens from reflecting off the viewing chamber walls onto the viewing and photographing area to cause "fog." The first two functions are achieved with accurately centered 75 micron platinum apertures. The apertures are identical and may be shifted from the objective to the collimating position as they become too contaminated for use in the objective gap. The small size of the aperture hole protects the specimen from damage from the intense electron beam and from radiant heat and evaporation products from the hot filament. The hole size is large enough to assure alignment of apertures without auxiliary mechanical positioning means and yet is small enough to give good contrast in the objective. The aperture following the projector is a 0.145 inch hole and is not critical in position or dimension.

ALIGNMENT

Alignment of the upper objective pole piece is a problem in vectors. Since the objective misalignment is manifested by an essentially linear shift of the image as the accelerating voltage is varied, the degree of misalignment may be measured by comparing shifts occurring for unit changes of voltage. It is found that the direction of image shift is different from the direction in which the center of the upper objective pole piece deviates from the true optical axis of the microscope. At 6000×, for medium misalignments, corresponding to less than 0.5 inch/1000 volts, the directions differ by about 120°, the misalignment direction being counterclockwise from the image shift when looking along the axis of the instrument toward the viewing screen. The magnitude of the change in image shift per kilovolt change in accelerating voltage is about 0.5 inch/1$K.V.$ for an 0.001 inch misalignment, or 75° turn on the yoke adjustment screws. Once an image shift has been noted as to direction and magnitude, the pole piece many be adjusted a compensating amount

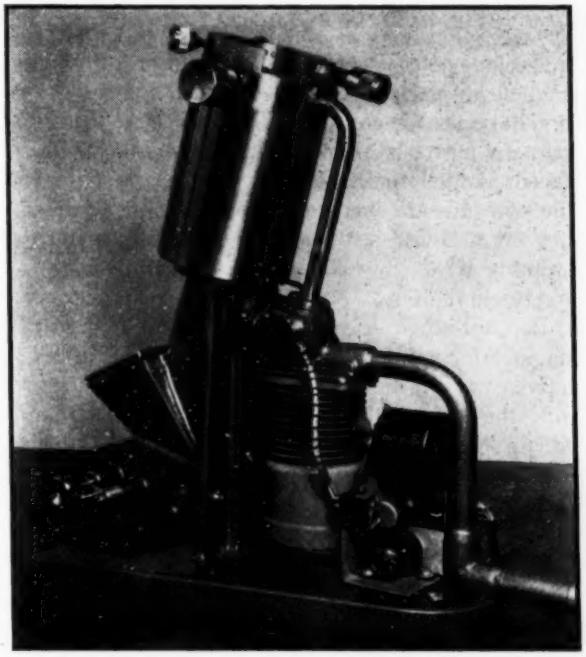


FIG. 6. A side view of the microscope with back cover removed to show arrangement of the vacuum system.

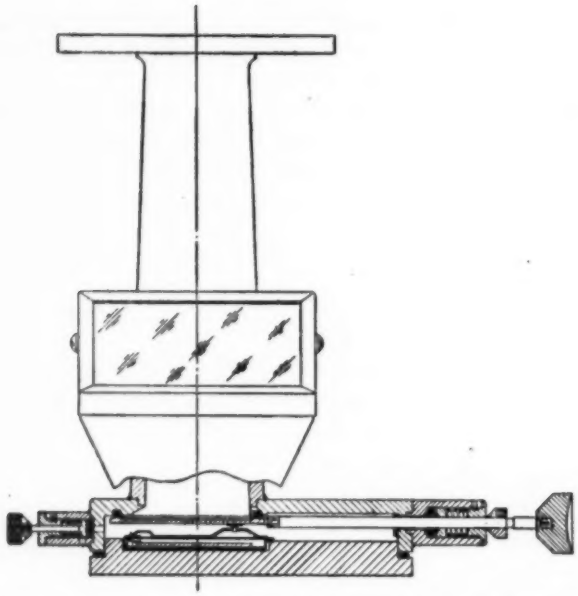


FIG. 7. Cross section showing the sliding value of the photographic chamber. The knob on the right is the activating means. The knob on the left admits air to the photographic chamber to permit opening the chamber door.

in the opposite direction 120° counterclockwise from that of the image shift. Usually, two or three attempts will reduce the image shift to a negligible value so that the only image change with voltage change is that of magnification. As previously stated, once the alignment yoke is adjusted and tightened, the alignment will hold while pole pieces are removed and changed through very long periods of use.

LENS ENERGIZATION

The lens pole piece gaps are energized by permanent magnets appropriately disposed in a magnetic circuit to reduce the deleterious effects of stray fields.* The magnetic means is Alnico V in four rectangular magnets placed parallel to the optical axis of the microscope. The two gaps are energized in parallel, but in the opposite sense, so that the outside shell of the magnetic assembly is sensibly a surface of constant mmf. This largely eliminates stray fields which can reduce magnification, produce asymmetries and distortions in the images, and make alignment difficult. The upper, objective gap gives a magnetomotive force of 1300 gilberts, while the lower, projector gap gives a magnetomotive force of 900 gilberts. These values are comparable with those from electromagnetic energizers of the same size. The magnetic circuit parameters are designed so as to give high flux leakage; i.e., to operate the magnets at a large demagnetization coefficient. Such design makes possible the frequent insertion and removal of pole pieces without fear of demagnetization of the energizing magnets. The magnets are energized after assembly. If the spool assembly is ever disassembled, the magnets will be demagnetized to the point where the lenses are not adequately energized.

In spite of the use of powerful magnets, it is not difficult to insert or remove pole piece elements from the bore of the magnetic spool assembly.

VACUUM SYSTEM

The volume of the entire microscope, excluding pumping elements, is approximately one liter. This volume is divided into two pumping circuits. The plate chamber and viewing chamber comprise about 70 percent of the evacuated space, and are pumped directly into the diffusion pump through a $1\frac{3}{8}$ inch opening. The specimen chamber and gun are pumped through a half-inch line disposed along the back of the instrument. The plate chamber accounts for eight percent, the gun 25 percent, and the specimen chamber three percent of the total volume. Figure 6 shows the vacuum system.

Two specialized valving means are provided to replace the versatile, but complex valve systems generally used on electron microscopes. Basically, only two reasons for breaking the vacuum exist in routine electron microscopy; the insertion or removal of specimens, and the insertion and removal of photographic plates. Simple, multi-purpose assembles have been devised to fulfill these specialized functions. The vacuum sealing surface of the specimen lock is also the bearing through which all specimen motion is transmitted into the chamber, while the valve plate which seals the photographic chamber also provides a fluorescent screen and serves as a photographic shutter.

Insertion of a specimen admits only 0.02 cc of air into the vacuum system, while withdrawal of a specimen admits no air. The pumping system will recover safe operating vacuum within six seconds after the admission of a specimen. The introduction of a photographic plate into the vacuum enclosure introduces an appreciable amount of air, 80 cc. This is a result of the necessity for providing space for a sliding valve plate and cassette cover during photographic exposures. Figure 7 shows the valving mechanism. The valve is a sliding door which is moved on and off a circular gasket cemented in the wall of the chamber casting. To pump a plate, the knob outside the chamber is pulled to slide the valve off its gasket. This action breaks the vacuum seal between the viewing chamber and the photographic chamber, permitting the latter to be pumped out. Returning the knob to its initial position seals off the photo-chamber. Air is admitted to the photo-chamber so that it may be opened by pulling a simple relief valve on the left side of the photo-chamber. The vacuum valve system is very dependable. Air may be admited to the photochamber after plate exposure, without turning off the electron beam. A time of about 60 seconds is required to recover operating vacuum after the introduction of a dry photographic plate.

All air introduced with specimens and plates must

* J. H. Reisner, J. App. Phys. **21**, 68 (1950). (In process of publication in full.)

pass through the diffusion pump. In spite of the deleterious effect which air has upon diffusion pump oil, octoil has been used for several months of general operation. With the much more stable silicone oil, very much longer periods of operation are possible.

When, in occasional maintenance, any of the components of the microscope are to be adjusted or removed from the vacuum, it is necessary to turn off the diffusion pump for several minutes before admitting air to the system through the relief valve on the photochamber.

The diffusion pump is a small air cooled D.P.I. VMF-10, using a Pressovac as a fore pump. The pump-down time from a cold start is about twenty-five minutes. A thermocouple gauge is used to indicate the vacuum. The meter employed is the test meter of the power supply which may be switched in when a pressure indication is desired.

SPECIMEN CONTROL

An unusual treatment of the specimen introduction and control has resulted in very great simplification in the operation and construction of the microscope. Grommet vacuum seals about shafts have long been used to obtain displacement and rotation through the confines of vacuum systems. By putting small cut-outs in such shafts, it is possible to introduce objects placed in the cut-out space into the vacuum, if the grommet is longer than the cut-out space. Such is the basis for the specimen holder to be described. Figures 4 and 5 show the actual specimen control system. A cylindrical hole 0.140 diameter, to accommodate the specimen holder, is open perpendicularly through a $\frac{5}{16}$ diameter shaft to bottom 0.012 inch from the surface of the shaft. A $\frac{1}{16}$ hole is continued through the shaft wall to permit egress of the electron beam. The transverse hole is roughly at the center of the six inch shaft. It is so positioned that when the shaft is pushed into the instrument up to the plunger of the right hand control, the specimen holder will be positioned over the lens opening. The area of the specimen is scanned in one direction by driving the shaft axially with the right-hand knob and plunger, and in a perpendicular direction by rotating with the left-hand knob.

The shaft rotation amounts to but 3° from its center position and permits scanning of three openings of a 200 mesh specimen support screen. The screen is situated 0.139 inch from the axis of rotation of the specimen shaft. In the extreme rotation positions, the specimen is only five microns closer to the objective lens than in its mean position. Such a displacement represents about 70 volts difference in focal voltages with the 3000× pole piece, and about 200 volts with the 6000× pole piece. The depth of focus is so great with the electron microscope that it is generally not necessary to refocus the field in going from the center to the extremes of specimen holder rotation. A total of about 36 mesh openings may be observed, 12 along the axis of the rod, three by rotation.

The axial movement of the specimen rod is provided through the right-hand knob by means of screw threads which advance a plunger registering against a hard metal plate on the end of the specimen holder rod. The rod is held tightly against the plunger by springs in the left-hand knob which grip a conical surface on the microscope housing. A limit spring slips into a groove at the right end of the specimen holder rod to prevent the rod from being pulled out of its sealing grommet and to prevent it from sliding back into the instrument under atmospheric pressure. The ability to remove the holder entirely from the path of the beam is useful in initially lining up the gun with the other optical elements.

Because of the sensitivity of the transverse control, six degrees of rotation to give full range of motion, the left-hand control knob is geared down 1:3. It has been found from experience that under these circumstances, control is quite adequate. The left-hand knob is highly damped in action to resist twist in the vacuum sealing grommet. In order that the specimen rod be correctly oriented to pass the electron beam through the specimen, a guide pin on the microscope must fit into a slot on the left hand knob before the rod can engage its springs on the advancing cone.

In order to keep the vacuum sealing grommet from wearing or binding, the specimen holder shaft is very lightly lubricated periodically with clear octoil.

VIEWING CHAMBER ASSEMBLY

In all electron microscopes, the viewing screen and plate must be normal to the beam. Direct viewing, desirable for brightness and sharpness, means that the fluorescent screen must be viewed obliquely. Therefore, the fifteen degree slope of this microscope column has been applied to facilitate direct viewing of the image. A wide window permits viewing with both eyes, and viewing by two people simultaneously.

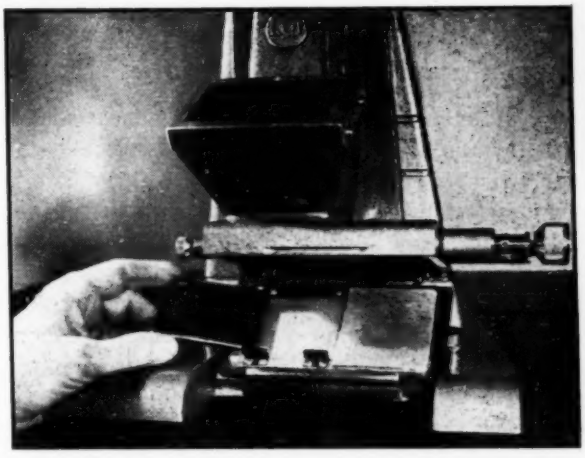


FIG. 8. Insertion of a photographic plate cassette.

Valving and plate exposure are accomplished by two moving parts, a rod with a hand knob on the end, and a metal plate. The plate has a circular depressed area on its upper surface to hold the electron sensitive phosphor and yet protect it from the vacuum sealing gasket across which the metal plate slides in the various phases of its valving cycle. A pin on the lower surface of the metal plate engages the cover to the photographic plate cassette. The description of a complete photographic cycle will make the operation of the mechanism more apparent.

With the camera door dropped open, a 2×2 plate cassette is clipped into an orienting slot and hold-down springs, Fig. 8. The door is snapped shut. The photochamber is pumped by sliding the movable plate far enough to the right to slip off the sealing gasket and open the chamber to the column. A stop on the shaft prevents more than the correct displacement. As soon as the photo chamber is evacuated, the high voltage may be turned on and an image focused. To expose the photographic plate, the same right hand knob is pulled to its extreme right position. This withdraws the valve plate from before the photographic plate. At the same time, a catch on the valve plate has engaged the cassette cover and slid it to the side, thus exposing the photographic plate. To end the exposure, the knob is pushed back to its inmost position, which closes the cassette cover and returns the valve plate to cover the vacuum gasket. A pull on a release valve on the left side of the photochamber admits air to the chamber which may then be opened for cassette removal.

The sliding valve plate is held against the vacuum gasket by light spring runners. The opening in the base of the viewing chamber is $1\frac{7}{8}$ inches in diameter, permitting exposure of a circular area of the same diameter on the photographic plate. The pole piece may be inserted and withdrawn through this hole with the aid of simple tools. The objective gap may be varied by engaging the screw slots on the projector by means of another simple tool inserted through the bottom hole when the valve plate is pulled aside.

POWER SUPPLY

Accelerating voltage and filament heating current are supplied by a fifty kilovolt supply almost identical with that used on commercial microscopes.[3] A driver unit which can sit on the table near the microscope permits control of voltage for focusing and control of emission for beam intensity. Once set with the variable controls, the chosen voltage and filament current reappear with every turn-on of the high voltage power switch. A meter reads, over its full scale, the reference voltage corresponding to the output range of voltages from 46 to 50 kv. Thus the focus voltage may be observed accurately on the meter. Such a feature is an excellent check on nearness to focus. It is useful in making through-focus series of photographs. It is a necessity in adjusting pole pieces to attain the desired focal voltage.

All electrical controls cause a voltage increase with a clockwise rotation. This means that the focal length of the lenses increases with a clockwise rotation of the high voltage focus control. As a result of the same electrical convention, electromagnetic lenses have their focal length decreased by the clockwise rotation. Using voltage focusing, the experienced electronmicroscopist finds the changes of appearance of fringes in the image occurring in reverse order to that which is encountered with current focusing. No other difference is apparent in the focusing process.

In an instrument designed for use by personnel untrained in electrical practice, safety precautions must be extensive. Circuit interlocks necessitate the correct positioning of the high voltage cable on the instrument and the attachment of the protective gun hood before voltage can be applied to the instrument. A quick-acting grounding mechanism is provided to discharge the high voltage supply when the gun hood is raised from its seat more than $\frac{3}{8}$ of an inch. If the high voltage has not been turned off before raising the hood, grounding the output in this manner will cause an overload relay system on the power supply to turn off the power.

A lead glass viewing window is provided to prevent the escape of x-rays generated by the electrons incident upon the viewing screen.

GENERAL FEATURES

The electron gun is the only externally adjustable optical element of the electron microscope. A heavy brass anode structure supports the filament grid and glass insulator components which are rigidly affixed to it. The brass anode is also the movable element, permitting sliding movement in a plane perpendicular to the optical axis of the microscope. The flat lower wall of the anode structure has holes to pass the electron beam and to permit evacuation of the gun volume through the specimen chamber. A gasket seat is provided on the bottom surface to carry the rubber gun gasket which provides a sliding vacuum seal between the anode and the specimen chamber. The position of the gun axis is determined by two adjustable plungers which register radially against the cylindrical outside wall of the anode. These plungers are activated by knobs, the two highest knobs on the column (Fig. 5), working through screw threads. A spring-loaded plunger extending radially backward on the microscope holds the anode tightly against the adjusting plungers to make the anode follow the plungers. The complete gun assembly may be removed from the column by removing the spring loaded plunger and sliding the gun to its extreme rear position so as to clear the adjusting plungers. If the position of the gun adjustment knobs is not changed, a gun may be removed and reinserted without changing alignment enough to lose illumination.

Since filaments, apertures, and diffusion pump oil

must be replaced occasionally and pole pieces must be removed for cleaning, all these components may be reached without major disassembly of the instrument or removal from its stand.

Stray alternating magnetic fields and vibration have always been two serious environmental hazards in high resolution electron microscopy. Careful consideration has been given toward overcoming these dangers in this instrument. From the objective pole piece to the photographic plate, the microscope beam is shielded behind heavy iron walls. In the extremely sensitive field region between lenses, the shielding from external fields is complete. Operation in the vicinity of relatively strong fields has shown little effect on the microscope image.

Vibration problems are largely eliminated by rigid construction throughout. Lens elements are held fast by the magnetic field in a one-piece housing. The viewing and specimen chambers are rigidly bolted to the lens housing. The two moving parts are the specimen rod and gun, both strong and well supported. The column of the instrument is supported at the base of the lens housing, not far from the center of gravity of the column, with the result that environmental vibration acts on the whole rather than on various parts.

SUMMARY

The instrument herein described has shown itself to be excellent for almost all types of electron microscopy not requiring resolutions better than 100A. Its unusual reproducibility makes it ideal for routine microscopy. Long sequences of pictures may be taken without touching gun adjustment, focus, or emission controls.

The direct magnifications of 1500, 3000, and 6000 times are chosen for convenience of observation and are not to be taken as a measure of resolution. Well-focused photographic images frequently reveal useful detail up to 20,000 diameters over-all magnification. The 1500× lens system is useful where studies are to be made on specimens familiar to the light microscopist and on specimens such as sections of tissue where fine structure must be viewed in relation to familiar gross structure such as cell walls. The 6000× magnification is useful where the highest resolution is routinely desired, such as in virus studies, and may most easily be achieved by virtue of the ability to focus the high magnification optical image most accurately. Figure 9 shows a high resolution photomicrograph taken with the described instrument.

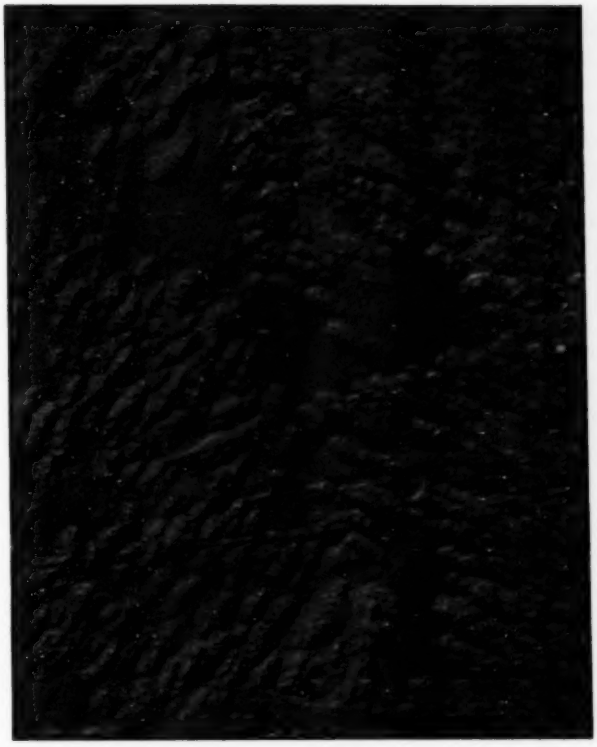


FIG. 9. Photomicrograph of a shadowed replica of an etched stainless steel surface. Magnification 20,000×. Taken at 3000×.

ACKNOWLEDGMENT

The authors wish to acknowledge the guidance and technical assistance of Perry C. Smith, under whose direction the work was carried out, and of Robert G. Picard, who frequently collaborated on the project. The electrical circuitry was carried out by Samuel M. Zollers. Acknowledgment is due Charles Yearsley, Richard McGarry, and A. M. Rennie for their assistance on mechanical design, and to Gustave Burger for much of the actual construction.

Artificial Field Equations for a Region Where μ and ϵ Vary with Position

P. D. P. SMITH
Westmount, Province of Quebec, Canada
(Received March 6, 1950)

I. Theory. The field equations developed admit of solutions in regions where μ and ϵ vary with position. For regions of zero conductivity, they are symmetrical with respect to electric and magnetic quantities, and the TM case can be derived from the TE. Fields derivable from a single component 4-potential are examined, and their propagation equations are found. A table relates the tensor and pseudo-tensor components to the usual vector components.

II. Applications in cylindrical polar coordinates. For the TE case, three types of solution exist, called A, B, or C, where the **E** field is parallel to the elementary vectors **dz**, **dr** or **rdθ**, respectively (for the TM case, **H** replaces **E**). In each TM or TE case, two possible distributions of μ and ϵ occur, leading to separable propagation equations. TEM waves can be derived from both TM and TE cases, but the impedances are different.

TE, TM and TEM waves for certain specific distributions of μ and ϵ, and their applications to curved guides, horns, and co-axial lines are considered. Some expressions are obtained for the fields, propagation constants, cut-off frequencies, and impedance. These solutions are not necessarily of the most general type, for to meet the conditions that restrict their choice, μ and ϵ are taken to be independent of the coordinate to which the transverse component of **E** (in TE) or **H** (in TM) is parallel.

I. THEORY

THE usual methods of procedure yield solutions of Maxwell's equations only in the following cases: (a) where space is everywhere homogeneous, (b) where space is subdivided into several homogeneous regions, (c) where the inductivity varies so slowly that it may be considered homogeneous.

This has been pointed out by Liebowitz,[1] who showed that it is possible to derive separable wave equations in a space where ϵ* is a function of position, by making use of a metric

$$ds^2=1/\epsilon\cdot\{(du^1)^2+(du^2)^2+(du^3)^2\}$$

and an auxiliary scalar function which is determined from ϵ. He proceeded to investigate the static field of a point charge. By assuming suitable distributions of ϵ, the field energy could be rendered finite.

However, a statement of the distribution of ϵ is a statement of the variation of the speed of light, and therefore of the variation of gravitational potential. It appears that his attack is really equivalent to guessing the solutions of Einstein's field equations for the space around a point charge. This particular instance does not concern the reduction in light velocity by scattering from many centers, as occurs in a dielectric, but a reduction in light velocity supposed inherent in the properties of space itself, such as that which causes the deviation of a light ray passing a heavy body. The problem of the gravitational field of a point charge has been solved by Weyl[2] and others, so that it is not necessary to guess at the result. The more fruitful use of equations of this type may be in problems of applied physics, to find solutions of Maxwell's equations for regions where μ and ϵ are known functions of position.

An alternative approach to Liebowitz's problem is formulated in this paper. The procedure used here is to set up a set of tensor field equations for a space where μ and ϵ are functions of position, and then investigate certain simple types of transverse electric and transverse magnetic waves, and the conditions for their propagation. The equations give results for regions where both μ and ϵ vary with position, and this respect are more general than Liebowitz's. On the other hand, solutions have only been found for certain simple cases where the four potential has only one component.

It turns out to be profitable to seek distributions of μ and ϵ leading to separable equations of propagation, rather than to seek general solutions for arbitrary μ and ϵ.

The field equations used are "artificial," in the sense that they are an artifice to help obtain solutions for the case of variable μ and ϵ. They have the appearance of being relativistic, but actually are only exact when the medium is at rest with respect to the observer. There is no practical disadvantage in this from an engineering point of view, as the velocity of the medium with respect to the observer is always far smaller than the velocity of light in the medium.

The equations, which have been published previously,[3] are

$$F^i=\partial_j F^{ij}, \tag{1.01}$$

$$0=\partial_i\varphi_{jk}+\partial_j\varphi_{ki}+\partial_k\varphi_{ij}, \tag{1.02}$$

$$F^{ij}=g^{\frac{1}{2}}\cdot(\epsilon/\mu)^{\frac{1}{2}}\varphi^{ij}, \tag{1.03}$$

$$F^i=g^{\frac{1}{2}}\rho\cdot dx^i/ds, \tag{1.04}$$

$$ds^2=g_{ij}dx^i dx^j=c^2dt^2/\mu\epsilon-\alpha^2(dx^1)^2-\beta^2(dx^2)^2-\gamma^2(dx^3)^2\simeq c^2dt^2/\mu\epsilon. \tag{1.05}$$

Equation (1.03) differs from the usual form because of the factor $(\epsilon/\mu)^{\frac{1}{2}}$. The current tensor, (1.04), becomes in an orthogonal coordinate system

$$F^i=\alpha\beta\gamma\rho\dot{x}^i/c(1-v^2/v_0^2)^{\frac{1}{2}},$$
$$v_0=c/(\mu\epsilon)^{\frac{1}{2}}.$$

[1] Liebowitz, Phys. Rev. **64**, 294 (1943).
* ϵ, the dielectric "constant."
[2] Weyl, *Space, Time, Matter* (Methuen and Company, Ltd., London, 1920), pp. 252–262.
[3] P. D. P. Smith, J. App. Phys. **20**, 633 (1949).

This shows that the expressions are only accurate when $v \ll v_0$, and are not truly relativistic.

Ohm's law, for this system, becomes

$$f_i = \sigma \varphi_{ai} \bar{u}^a, \quad i<4, \tag{1.06}$$

where $\bar{u}^a$ is the mean velocity defined by

$$\bar{u}^a = \int u^a \cdot dN \Big/ \int dN,$$

$\int dN$ being the number of electrons in a volume that is small enough so that g_{ij} may be considered a constant, σ representing the conductivity. The Eq. (1.01) becomes

$$\{\sigma \cdot (\mu/\epsilon)^{\frac{1}{2}} \bar{u}_j + \partial_j\} F^{ij} = 0, \quad i<4 \tag{1.07}$$

equivalent to

$$\sigma\{\mathbf{E} + (1/c)[\bar{\mathbf{v}} \times \mathbf{B}]\} = \text{curl}\mathbf{H} - \dot{\mathbf{D}}/c \tag{1.08}$$

in vector notation.

In regions where $\sigma = 0$, there is complete symmetry between the electric and magnetic fields when the roles of μ and ϵ are interchanged. This symmetry can be brought out by introducing a new four-potential, ψ_i, the counterpart of the usual φ_i. We had

$$\varphi_{ij} = \partial_i \varphi_j - \partial_j \varphi_i \tag{1.09}$$

from which can be defined the pseudo-tensor

$$B^{ij} = (-)^{p(i,\, j,\, k,\, l)} \varphi_{kl}, \tag{1.10}$$

where $p(i, j, k, l)$ is the number of inversions of i, j, k, l from the order 1, 2, 3, 4. Thus it follows

$$F^{ij} = (g\epsilon/\mu)^{\frac{1}{2}} g^{ia} g^{jb} (-)^{p(a,\, b,\, c,\, d)} B^{cd} \tag{1.11}$$

and conversely

$$B^{ij} = (\mu/g\epsilon)^{\frac{1}{2}} g_{ka} g_{lb} (-)^{p(i,\, j,\, k,\, l)} F^{ab}. \tag{1.12}$$

At this point the second type of field will be introduced, namely that produced by the alternative four-potential ψ_i. For these fields

$$\psi_{cd} = \partial_c \psi_d - \partial_d \psi_c, \tag{1.13}$$

$$\partial_b \psi_{cd} + \partial_c \psi_{db} + \partial_d \psi_{bc} = 0 \tag{1.14}$$

and

$$F^{ab} = (-)^{p(a,\, b,\, c,\, d)} \psi_{cd}. \tag{1.15}$$

Using this in (1.12),

$$B^{ij} = (-)^{p(i,\, j,\, k,\, l) + p(a,\, b,\, c,\, d)} (\mu/g\epsilon)^{\frac{1}{2}} g_{ka} g_{lb} \psi_{cd}. \tag{1.16}$$

The majority of practical problems will be set up in orthogonal coordinates, in which (1.16) simplifies to

$$B^{ij} = (-)^{p(i,\, j,\, k,\, l) + p(k,\, l,\, i,\, j)} \cdot (\mu/g\epsilon)^{\frac{1}{2}} g_{kk} g_{ll} \psi_{ij}$$

$$= (\mu/g\epsilon)^{\frac{1}{2}} g_{kk} g_{ll} \psi_{ij}. \tag{1.17}$$

Thus in orthogonal coordinate systems

$$B^{ij} = (\mu/\epsilon)^{\frac{1}{2}} (g_{kk} g_{ll} / g_{ii} g_{jj})^{\frac{1}{2}} \psi_{ij} \tag{1.18}$$

TABLE I. The co-variant and contra-variant fields generated from the four potentials φ_i and ψ_i, and their equivalents in vector notation. The components generated by φ_3 or ψ_3 are indicated by an asterisk.

TYPE 1 FIELD 4 POTENTIAL φ_i	TYPE 2 FIELD 4 POTENTIAL ψ_i
$\varphi_{41} = i\alpha E_1$	$\psi_{41} = -\alpha H_1$
$\varphi_{42} = i\beta E_2$	$\psi_{42} = -\beta H_2$
$\varphi_{43} = i\gamma E_3$ *	$\psi_{43} = -\gamma H_3$ *
$F^{41} = i\beta\gamma D_1$	$B^{41} = -\beta\gamma B_1$
$F^{42} = i\gamma\alpha D_2$	$B^{42} = -\gamma\alpha B_2$
$F^{43} = i\alpha\beta D_3$ *	$B^{43} = -\alpha\beta B_3$ *
$F^{12} = \gamma H_3$	$B^{12} = -i\gamma E_3$
$F^{23} = \alpha H_1$ *	$B^{23} = -i\alpha E_1$ *
$F^{31} = \beta H_2$ *	$B^{31} = -i\beta E_2$ *
$\varphi_{12} = \alpha\beta B_3$	$\psi_{12} = -i\alpha\beta D_3$
$\varphi_{23} = \beta\gamma B_1$ *	$\psi_{23} = -i\beta\gamma D_1$ *
$\varphi_{31} = \gamma\alpha B_2$ *	$\psi_{31} = -i\gamma\alpha D_2$ *

and corresponding to this in the first type field

$$F^{ij} = (\epsilon/\mu)^{\frac{1}{2}} (g_{kk} g_{ll} / g_{ii} g_{jj})^{\frac{1}{2}} \varphi_{ij}. \tag{1.19}$$

From (1.18) and (1.19) is built the following table connecting the tensor and pseudo-tensor components with the usual vector components. As before[3] $\varphi_{41} = i\alpha E_1$, $\varphi_{12} = \alpha\beta B_3$, and use is made Eqs. (1.03), (1.10), and (1.15).

Thus corresponding to any solution of the field equations generated by the 4-potential φ_i there can be found another solution generated by the alternative potential ψ_i. This use of two alternative solutions is very common in wave guide theory, one solution yielding transverse electric, or TE, waves, the other transverse magnetic, or TM, waves.

The TE Case

For the sake of simplicity, the examples solved will be restricted to the case where there is only one non-vanishing component of the 4-potential, so that

$$\varphi_i = \delta_{i3} \varphi. \tag{1.20}$$

The field equations give immediately

$$\begin{aligned} &\varphi_{12} = 0, && \varphi_{41} = 0, \\ &\varphi_{23} = \partial_2 \varphi, && \varphi_{42} = 0, \\ &\varphi_{31} = -\partial_1 \varphi, && \varphi_{43} = \partial_4 \varphi. \end{aligned} \tag{1.21}$$

The contravariant field components are, from (1.19) and (1.21),

$$\begin{aligned} &F^{12} = 0, && F^{41} = 0, \\ &F^{23} = \alpha/\beta\gamma\mu \cdot \partial_2 \varphi, && F^{42} = 0, \\ &F^{31} = -\beta/\gamma\alpha\mu \cdot \partial_1 \varphi, && F^{43} = (\epsilon\alpha\beta/\gamma)\partial_4 \varphi. \end{aligned} \tag{1.22}$$

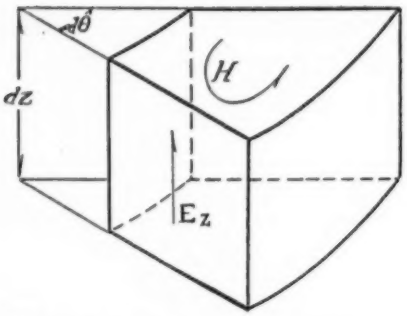


FIG. 1. TE waves of Type A. Fields generated by a single component 4-potential. The electric field is parallel to the z-axis, propagation is radial, or perpendicular to the axial planes.

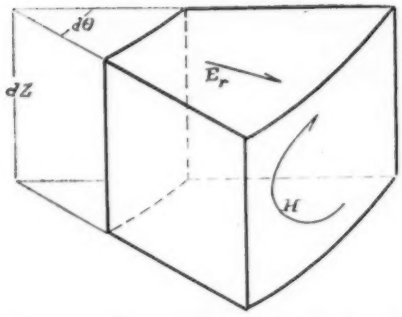


FIG. 2. TE waves of Type B. Fields generated by a single component 4-potential. The electric field is radial, and propagation is along the z-axis, or perpendicular to axial planes.

These non-vanishing components arising from φ_3, ψ_3 are marked with an asterisk in Table I, when their vector counterparts may be read.

Thus the contravariant group of field equations

$$0=\partial_j F^{ij}$$

become

$$0=\partial_1(\beta/\gamma\alpha\mu\cdot\partial_1\varphi)+\partial_2(\alpha/\beta\gamma\mu\cdot\partial_2\varphi)+\partial_4(\epsilon\alpha\beta/\gamma\cdot\partial_4\varphi)=0 \quad (1.23)$$

and also the three equations

$$\partial_3 F^{i3}=0. \quad (1.24)$$

These equations obviously give a TE wave propagated in directions perpendicular to dx^3. The electric field is parallel to the third coordinate dx^3.

If α, β, γ are independent of x^3, the equations (1.24) are identically satisfied if μ, ϵ, and φ are also independent of x^3. This restriction has been used in all cases examined in this paper.

On the other hand, it is simple to eliminate the field quantities from Eqs. (1.24), obtaining a pair of differential equations involving μ and ϵ. These, included merely for completeness, are

$$\partial_3 \log\{\partial_1 X/\partial_3(n^2\alpha^2)\}=0, \quad (1.25)$$

$$\partial_3 \log\{\partial_2 X/\partial_3(n^2\beta^2)\}=0, \quad (1.26)$$

$$n^2=\mu\epsilon, \quad X=\partial_3 \log\{\gamma/\epsilon\alpha\beta\}. \quad (1.27)$$

The TM Case

The TM case is obtained immediately from the TE case by interchanging the roles of μ and ϵ, ψ and φ. There follows, for the non-vanishing field components

$$\psi_{23}=\partial_2\psi, \quad \psi_{31}=-\partial_1\psi, \quad \psi_{43}=\partial_4\psi, \quad B^{23}=\alpha/\beta\gamma\epsilon\cdot\partial_2\psi, \quad (1.28)$$

$$B^{31}=-\beta/\gamma\alpha\epsilon\cdot\partial_1\psi, \quad B^{43}=\mu\alpha\beta/\gamma\cdot\partial_4\psi. \quad (1.29)$$

The propagation equation becomes

$$0=\partial_1(\beta/\gamma\alpha\epsilon\cdot\partial_1\psi)+\partial_2(\alpha/\beta\gamma\epsilon\cdot\partial_2\psi)+\partial_4(\mu\alpha\beta/\gamma\cdot\partial_4\psi) \quad (1.30)$$

and also

$$\partial_3 B^{i3}=0. \quad (1.31)$$

As before, 1.31 are satisfied if α, β, γ, μ, ϵ, and ψ are independent of x^3. If this is not so, then it would be necessary to ensure that μ and ϵ satisfy

$$\partial_3 \log\{\partial_1 Y/\partial_3(n^2\alpha^2)\}=0,$$
$$\partial_3 \log\{\partial_2 Y/\partial_3(n^2\beta^2)\}=0,$$

where

$$Y=\partial_3 \log(\gamma/\mu\alpha\beta).$$

This completes the theoretical part of this work. The next part involves the application of the theory to certain problems in cylindrical polar coordinates.

II. SOLUTIONS IN CYLINDRICAL POLARS

1. General Properties in Cylindrical Coordinates

Before proceding to consider specific examples using the theory developed in Section I, the general nature of the solutions which can arise from a single component of potential will be considered. The TM case has to be restricted to coordinate systems in which x^1 constant, x^2 constant, is a closed curve, because there is only one B component, and the corresponding lines of flux must be closed curves.

The possible choices are shown in Table II and Table III, along with the identifications of dx^1, dx^2, dx^3, α, β, and γ, and the corresponding propagation equation. The nature of the fields is shown diagrammatically in Figs. 1, 2, and 3 for TE waves, Fig. 4 for TM waves.

The fields are classified as Type A, B, or C, according as the coordinates (dx^1, dx^2, dx^3) are taken as (dr, $d\theta$, dz), ($d\theta$, dz, dr) or (dz, dr, $d\theta$).

The non-vanishing field components can be verified by referring to Table I of Section I, and remembering that only the third component of potential, φ_3 or ψ_3, is to be considered.

The form of TM waves conjugate to the TE waves of type C are shown in Table III. The equations $\partial_3 B^{i3}$ are automatically satisfied if ψ, μ, ϵ are independent of θ, since α, β, and γ are independent of θ.

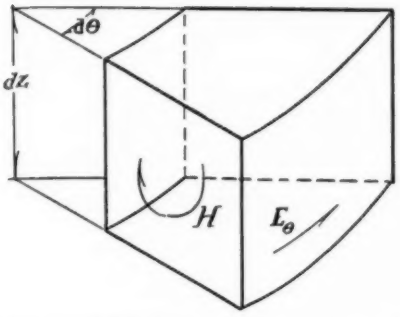


FIG. 3. TE waves of Type C. Fields generated by single component 4-potential. The electric field is perpendicular to planes through the axis, and propagation is radial or along the axis.

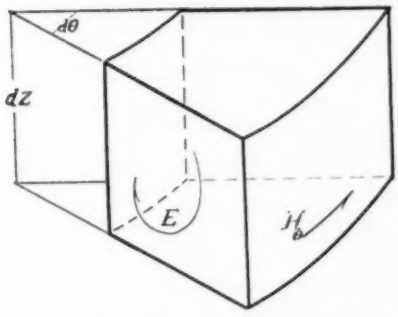


FIG. 4. TM waves of Type C. Fields generated by a single component of the alternative 4-potential. The magnetic field is perpendicular to planes through the axis, and propagation is radial, or along the axis.

TEM Waves

In addition to TE and TM waves, TEM waves can be propagated also. These are characterized by the potential depending on only one coordinate variable. For example, in the TM case the electric fields are given by

$$-i\beta\gamma\cdot D_1=\partial_2\psi, \quad -i\gamma\alpha\cdot D_2=-\partial_1\psi. \tag{2.01}$$

For TEM waves in a co-axial line, $D_z=0$, and hence it is necessary that $\partial\psi/\partial r=0$. The propagation equation reduces to

$$\frac{\partial}{\partial z}\left\{1/\epsilon\cdot\frac{\partial\psi}{\partial z}\right\}+\omega^2/c^2\mu\psi=0. \tag{2.02}$$

The field components are then given by

$$D_r=-i/r\cdot\partial\psi/\partial z, \quad H_\theta=-\omega/cr\cdot\psi. \tag{2.03}$$

The ratio

$$Z=E_r/H_\theta=ic/\omega\epsilon\cdot\partial/\partial z\log\psi. \tag{2.04}$$

The TE fields of type B, with $B_z=0$, can also yield a TEM wave. If $B_z=0$, then referring to the propagation equation for type B waves on Table II, in this case, where

$$B_z=-1/r\cdot\partial\varphi/\partial\theta=0, \tag{2.05}$$

it becomes

$$\partial/\partial z\{1/\mu\cdot\partial/\partial z(r\varphi)\}+\omega^2/c^2\cdot\epsilon\cdot r\varphi=0. \tag{2.06}$$

The field impedance ratio in this case becomes

$$Z'=E_r/H_\theta=(\omega\mu/ic)/\partial/\partial z\log\varphi. \tag{2.07}$$

If ψ and φ are proportional to $e^{S\text{ or }T}$, then for the TEM waves derived from TM waves

$$Z=\{\mu/\epsilon+c^2/\omega^2\epsilon^2\cdot(S''-S'\epsilon'/\epsilon)\}^{\frac{1}{2}}, \tag{2.08(a)}$$

while for those derived from TE waves

$$Z'=\{\epsilon/\mu+c^2/\omega^2\mu^2(T''-T'\mu'/\mu)\}^{-\frac{1}{2}}. \tag{2.08(b)}$$

When μ and ϵ are constant,

$$Z=Z'.$$

The next sections deal with specific examples of the method in greater detail. Only a few of the many possible examples have been solved.

Section 2. TE Waves of Type A

In this case

$$dx'=dr, \quad dx^2=d\theta, \quad dx^3=dz, \quad \alpha=1, \quad \beta=r, \quad \gamma=1. \tag{2.09}$$

The non-vanishing field components are

$$E_z=\omega/ic\cdot\varphi,$$
$$B_r=1/r\cdot\partial\varphi/\partial\theta, \quad B_\theta=-\partial\varphi/\partial r. \tag{2.10}$$

The propagation equation becomes, in this case

$$0=\frac{\partial}{\partial r}\left\{\frac{r}{\mu}\frac{\partial\varphi}{\partial r}\right\}+\frac{\partial}{\partial\theta}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial\theta}\right\}+\frac{\omega^2}{c^2}\epsilon r\varphi. \tag{2.11}$$

The remaining equations,

$$\partial/\partial z(F^{23})=0 \tag{2.12}$$

are automatically satisfied if φ, μ, and ϵ are independent of z, because α, β, γ are all independent of z.

The next problem is to obtain a possible pair of functions

$$\mu=\mu(r,\theta), \quad \epsilon=\epsilon(r,\theta)$$

that make (2.11) separable. Let

$$\mu=a(r)b(\theta), \quad \epsilon=c(r)f(\theta) \tag{2.13}$$

and let

$$\varphi=R(r)\Theta(\theta).$$

Then Eq. (2.11) becomes

$$\frac{ra(r)}{R}\frac{d}{dr}\left\{\frac{r}{a(r)}\frac{dR}{dr}\right\}+\frac{b(\theta)}{\Theta}\frac{d}{d\theta}\left\{\frac{1}{b(\theta)}\frac{d\Theta}{d\theta}\right\}$$
$$+\omega^2/c^2\cdot r^2a(r)c(r)b(\theta)f(\theta)=0.$$

This equation is separable if either

$$r^2a(r)c(r)=k_r \text{ a constant, or,} \tag{2.14}$$

$$b(\theta)f(\theta)=k_\theta \text{ a constant.} \tag{2.15}$$

These two cases will be referred to as k_r-separability and k_θ-separability.

TABLE II. Propagation equations and field components for TE waves generated by a single component of 4-potential.

dx^1	dx^2	dx^3	α	β	γ	Propagation Equation + Field Components	Character of Waves
dr	$d\theta$	dz	1	r	1	$\frac{\partial}{\partial r}\left\{\frac{r}{\mu}\frac{\partial\varphi}{\partial r}\right\}+\frac{\partial}{\partial\theta}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial\theta}\right\}+\frac{\omega^2}{c^2}\epsilon r\varphi=0$	(a) TE Waves in a curved guide travelling in direction $d\theta$. (b) TE Waves in a horn, travelling in direction dr. (c) TE Waves travelling radially in a cavity with cylindrical walls.
	TYPE A					$E_z=\omega/ic\,\varphi$, $B_r=1/r\cdot\frac{\partial\varphi}{\partial\theta}$, $B_\theta=-\frac{\partial\varphi}{\partial r}$	
$d\theta$	dz	dr	r	1	1	$\frac{\partial}{\partial\theta}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial\theta}\right\}+\frac{\partial}{\partial z}\left\{\frac{r}{\mu}\frac{\partial\varphi}{\partial z}\right\}+\frac{\omega^2}{c^2}\epsilon r\varphi=0$	(a) TE Waves in a guide with inclined walls, travelling in direction dz. (b) TE Waves in a curved guide, travelling in direction $d\theta$.
	TYPE B					$E_r=\omega/ic\,\varphi$, $B_\theta=\frac{\partial\varphi}{\partial z}$, $B_z=-1/r\frac{\partial\varphi}{\partial\theta}$	
dz	dr	$d\theta$	1	1	r	$\frac{\partial}{\partial z}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial z}\right\}+\frac{\partial}{\partial r}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial r}\right\}+\frac{\omega^2}{c^2}\frac{\epsilon}{r}\varphi=0$	(a) TE Waves in a guide with inclined walls travelling in a direction dz. (b) TE Waves in a horn, travelling in a direction dr.
	TYPE C					$E_\theta=\omega/ic\cdot\varphi$, $B_z=1/r\cdot\frac{\partial\varphi}{\partial r}$, $B_r=-1/r\cdot\frac{\partial\varphi}{\partial z}$	

For k_r-separability

$$\frac{ra(r)}{R}\frac{d}{dr}\left\{\frac{r}{a(r)}\frac{dR}{dr}\right\}=\lambda^2, \tag{2.16}$$

$$0=\frac{b(\theta)}{\Theta}\frac{d}{d\theta}\left\{\frac{1}{b(\theta)}\frac{d\Theta}{d\theta}\right\}+\{\lambda^2+\omega^2/c^2\cdot k_r b(\theta)f(\theta)\}. \tag{2.17}$$

On the other hand, k_θ-separability gives

$$\frac{ra(r)}{R}\frac{d}{dr}\left\{\frac{r}{a(r)}\frac{dR}{dr}\right\}=\{\eta^2-\omega^2/c^2\cdot k_\theta\cdot r^2a(r)c(r)\}, \tag{2.18}$$

$$\frac{b(\theta)}{\Theta}\frac{d}{d\theta}\left\{\frac{1}{b(\theta)}\frac{d\Theta}{d\theta}\right\}+\eta^2=0. \tag{2.19}$$

An Example of TE Waves of Type A Using k_θ-Separability

Either of these last two alternatives can be attacked as a new differential equation, the values of μ and ϵ being given. An alternative attack is to choose the values of $a(r)$ and $c(r)$ in such a way as to generate known differential equations, or, what is equivalent, to choose the functions R and Θ and seek as solutions $a(r)$ and $c(r)$. This provides a great flexibility for design of apparatus. The function $b(\theta)$ must be periodic if θ can vary over a range greater than 2π; for example, it could be

$$b(\theta)=1+A\cos k\theta$$

or it could even be represented by a Fourier series.

Equation (2.19) can also be written

$$d^2\Theta/d\theta^2+d\Theta/d\theta\cdot d/d\theta(\log 1/b(\theta))+\eta^2\Theta=0. \tag{2.20}$$

For the case where $b(\theta)=\theta^{-n}$, the solution is

$$\Theta=\theta^{\frac{1}{2}(1-n)}\{CJ_{\frac{1}{2}(n-1)}(\eta\theta)+DN_{\frac{1}{2}(n-1)}(\eta\theta)\}. \tag{2.21}$$

As materials of zero or infinite magnetic permeability are unknown, the plane $\theta=0$ must be outside the region filled with this particular medium. This could occur in a V-shaped horn with metal walls lying in the planes $\theta=\theta_1$, $\theta=\theta_2$, where

$$0<\theta_1<\theta_2<2\pi.$$

In such a horn η will be determined from the boundary condition that

$$E_z=\omega/ic\cdot\varphi=0$$

at

$$\theta=\theta_1 \quad \text{and} \quad \theta=\theta_2.$$

From this it follows that for a horn the boundary conditions are

$$J_{\frac{1}{2}(n-1)}(\eta\theta_1)/N_{\frac{1}{2}(n-1)}(\eta\theta_1)=J_{\frac{1}{2}(n-1)}(\eta\theta_2)/N_{\frac{1}{2}(n-1)}(\eta\theta_2). \tag{2.22}$$

The differential equation for the radial factor can be written from (2.18)

$$0=R''+R'(1/r-a'/a)+\{\omega^2/c^2\cdot k_\theta a(r)c(r)-\eta^2/r^2\}R. \tag{2.23}$$

At some radius

$$r_0 \text{ let } a_0=a(r_0), \quad c_0=c(r_0) \tag{2.24}$$

and let

$$v=\omega/c\cdot(k_\theta a_0c_0)^{\frac{1}{2}}\cdot r=\omega/c\cdot(\mu_0\epsilon_0)^{\frac{1}{2}}r. \tag{2.25}$$

Then we find

$$0=\frac{d^2R}{dv^2}+\frac{dR}{dv}\cdot\left\{\frac{1}{v}-\frac{d}{dv}\log a\right\}+\left\{\frac{a(r)c(r)}{a_0c_0}-\frac{\eta^2}{v^2}\right\}R. \tag{2.26}$$

TABLE III. Propagation equations and field components for TM waves generated by a single component of 4-potential.

dx^1	dx^2	dx^3	α	β	γ	Propagation Equation + Field Components	Character of Waves.
dz	dr	$d\theta$	1	1	τ	$\frac{\partial}{\partial z}\left\{\frac{1}{\tau\epsilon}\frac{\partial \Psi}{\partial z}\right\}+\frac{\partial}{\partial r}\left\{\frac{1}{\tau\epsilon}\frac{\partial \Psi}{\partial r}\right\}+\frac{\omega^2\mu}{c^2\tau}\Psi=0$	(a) TM waves propagated in direction dz in a cylindrical guide or co-axial line.
TM WAVES TYPE C						$H_\theta=-\omega/c\tau\cdot\Psi$ $\quad D_z=i/\tau\cdot\frac{\partial \Psi}{\partial r}$ $\quad D_r=-i/\tau\frac{\partial \Psi}{\partial z}$	(b) TM waves propagated radially in a 360° horn. (c) TM Waves propagated radially in a cavity with cylindrical walls.

Comparing this with the last equation given by Janke and Emde in their Chapter VIII,[4] Section 7*, which is, with x replaced by v

$$\frac{d^2y}{dv^2}+\frac{dy}{dv}\cdot\left(\frac{1}{v}-2u\right)+\left\{1+u^2-\frac{du}{dv}-\frac{u}{v}-\frac{p^2}{v^2}\right\}y=0 \quad (2.27)$$

it is clear the two Eqs. (2.26) and (2.27) are identical if

$$2u=d/dv\cdot\log a=(a'/a)\cdot c/\omega(\mu_0\epsilon_0)^{\frac{1}{2}}, \quad (2.28)$$

$$1+u^2-\frac{du}{dv}-\frac{u}{v}=\frac{a(r)c(r)}{a_0c_0}=\frac{a(v)c(v)}{a_0c_0}, \quad (2.29)$$

$$p^2=\eta^2. \quad (2.30)$$

Thus the solution becomes

$$R=a(r)^{\frac{1}{2}}\{CJ_\eta(\omega/c\cdot(\mu_0\epsilon_0)^{\frac{1}{2}}\cdot r)+DN_\eta(\omega/c\cdot(\mu_0\epsilon_0)^{\frac{1}{2}}r). \quad (2.31)$$

The function $c(v)=c(r)$

$$=\frac{a_0c_0}{a(r)}\left\{1-\frac{c^2}{2\omega^2\mu_0\epsilon_0}\left(\frac{a''}{a}-\frac{3}{2}\left(\frac{a'}{a}\right)^2+\frac{a'}{ra}\right)\right\}. \quad (2.32)$$

Thus we have, collecting results, for the case where this medium fills a conducting horn from $\theta=\theta_1$ to $\theta=\theta_2$,

$$\mu=a(r)\cdot\theta^{-n}, \quad (2.33)$$

$$\epsilon=(\mu_0\epsilon_0\theta^n/a(r))\{1-c^2/2\omega^2\mu_0\epsilon_0\cdot(a''/a-\tfrac{3}{2}(a'/a)^2+a'/ra)\}, \quad (2.34)$$

$$\varphi=\theta^{-q}\{J_q(\eta\theta)/J_q(\eta\theta_1)-N_q(\eta\theta)/N_q(\eta\theta_1)\}\times a(r)^{\frac{1}{2}}\cdot H_\eta^{(1)\text{ or }(2)}(\omega(\mu_0\epsilon_0)^{\frac{1}{2}}r/c),$$

$$E_z=\omega/ic\cdot\varphi, \quad B_r=1/r\partial\varphi/\partial\theta, \quad B_\theta=-\partial\varphi/\partial r, \quad (2.10)$$

where

$$q=\tfrac{1}{2}(n-1).$$

These solutions contain an arbitrary function $a(r)$.

[4] Janke and Emde, *Tables of Functions* (Dover Publications, New York, 1943), p. 147.

An Example of TE Waves of Type A Using k_r-Separability

In the case of k_r-separability, the differential equation for R is, from (2.16),

$$R''+(1/r-a'/a)R'-\lambda^2/r^2\cdot R=0. \quad (2.35)$$

This becomes a very simple equation, if the medium is constructed so that

$$a(r)=Cr^A, \quad (2.36)$$

$$a'/a=A/r, \quad (2.37)$$

for then R satisfies the homogeneous equation

$$R''+(1-A)R'/r-\lambda^2R/r^2=0, \quad (2.38)$$

so that the solution is $R=Pr^{m_1}+Qr^{m_2}$; with m_1 and m_2 the two roots of $m(m-1)+m(1-A)-\lambda^2=0$ or

$$m=\tfrac{1}{2}A\pm(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}. \quad (2.39)$$

Thus when $a(r)=Cr^A$ $c(r)=k_rC^{-1}r^{-A-2}$ the factor R is

$$R=r^{\frac{1}{2}A}\{P\exp[\log r(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}]+Q\exp[-\log r(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}]\}. \quad (2.40)$$

The Θ factor satisfies (2.17), or

$$0=\Theta''+\Theta'\cdot d/d\theta\cdot\log 1/b(\theta)+\{\lambda^2+\omega^2k_rf(\theta)b(\theta)/c^2\}\Theta. \quad (2.41)$$

Again, considering the case where $b(\theta)=\theta^{-n}$, the equation becomes

$$\Theta''+n\Theta'/\theta+\{\lambda^2+\omega^2k_rf(\theta)/c^2\theta^n\}\Theta=0. \quad (2.42)$$

As in the previous section, the factor is taken as

$$\Theta=\theta^{\frac{1}{2}(1-n)}J(\theta). \quad (2.43)$$

J then satisfies

$$0=J''+J'/\theta+\{\lambda^2+\omega^2k_rf(\theta)/c^2\theta_n-\tfrac{1}{4}(1-n)^2/\theta^2\}J. \quad (2.44)$$

The equation evidently becomes a Bessel's equation if

$$f(\theta)=L^2\theta^n+\theta^{n-2}$$

the solutions being

$$\Theta=\theta^{\frac{1}{2}(1-n)}H_p^{(1)\text{ or }(2)}(\lambda_0\theta), \tag{2.45}$$

where $p^2=\frac{1}{4}(n-1)^2-\omega^2k_r/c^2$, p and λ_0 are constants, $\lambda_0^2=\lambda^2+\omega^2k_rL^2/c^2$.

Summarizing the results of this example, in a medium where

$$\mu=Cr^A\theta^{-n}, \tag{2.46}$$

$$\epsilon=k_rC^{-1}r^{-A-2}(L^2\theta^n+\theta^{n-2}) \tag{2.47}$$

the function φ is

$$\varphi=r^{\frac{1}{2}A}\{P\exp[\log r(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}] +Q\exp[-\log r(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}]\} \times\theta^{\frac{1}{2}(1-n)}\{EH_p^{(1)}(\lambda_0\theta)+FH_p^{(2)}(\lambda_0\theta)\}, \tag{2.48}$$

$$p^2=\tfrac{1}{4}(n-1)^2-\omega^2k_r/c^2, \quad \lambda_0^2=\lambda^2+L^2\omega^2k_r/c^2. \tag{2.49}$$

For a more specific example, a curved wave guide, bounded by highly conducting metal sheets at two planes of constant z, and by two cylinders of radii r_0 and r_1, may be considered. The variable r may be written

$$r=e^{\sigma}$$

and to make $\varphi(r_0)=0$ and $\varphi(r_1)=0$ (so that E_z vanishes at its outer and inner wall), let

$$P=e^{i\alpha_0},$$
$$Q=-e^{-i\alpha_0}.$$

Then in this case

$$R=r^{\frac{1}{2}A}\{\exp[\sigma(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}+i\alpha_0] -\exp[-\sigma(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}-i\alpha_0]. \tag{2.50}$$

This can have zeros for real values of r or σ if

$$\lambda^2+\tfrac{1}{4}A^2<0.$$

Writing

$$\lambda^2=-m^2,$$
$$(\lambda^2+\tfrac{1}{4}A^2)^{\frac{1}{2}}=i(m^2-\tfrac{1}{4}A^2)^{\frac{1}{2}}$$

and this makes

$$R=r^{\frac{1}{2}A}\cdot 2i\cdot\sin\{(m^2-\tfrac{1}{4}A^2)^{\frac{1}{2}}\sigma+\alpha_0\}. \tag{2.51}$$

When $r=r_0$, $\sigma=\sigma_0$, and $r=r_1$, $\sigma=\sigma_1$, $R=0$, so

$$(m^2-\tfrac{1}{4}A^2)^{\frac{1}{2}}\sigma_0+\alpha_0=0,$$
$$(m^2-\tfrac{1}{4}A^2)^{\frac{1}{2}}\sigma_1+\alpha_0=l\pi, \quad l=\text{integer}$$
$$m^2-\tfrac{1}{4}A^2=l^2\pi^2/(\sigma_1-\sigma_0)^2.$$

Since the Hankel's functions contain exponential factors of the type $\exp(\pm i\lambda_0\theta)$, in this case they become proportional to

$$\exp\pm i\{\omega^2k_rL^2/c^2-l^2\pi^2/(\sigma_1-\sigma_0)^2-\tfrac{1}{4}A^2\}^{\frac{1}{2}}\cdot\theta.$$

This guide will pass all frequencies above f_c where

$$4\pi^2f_c^2k_rL^2/c^2>l^2\pi^2/(\sigma_1-\sigma_0)^2+\tfrac{1}{4}A^2.$$

It will be noted f_c depends upon the two constants A and L.

Section 3. TE Fields of Type C

The specific case considered will be one whose

$$\mu=Mr^{2k-2}e^{2\alpha z}, \tag{3.01}$$

$$\epsilon=\epsilon_0\mu_0/\mu. \tag{3.02}$$

For fields of type C

$$dx^1=dz, \quad dx^2=dr, \quad dx^3=d\theta,$$
$$\alpha=1, \quad \beta=1, \quad \gamma=r. \tag{3.03}$$

The non-vanishing field components are

$$E_\theta=\omega/icr\cdot\varphi,\ B_z=1/r\cdot\partial\varphi/\partial r,\ B_r=-1/r\cdot\partial\varphi/\partial z. \tag{3.04}$$

The propagation equation is, from Table II,

$$\frac{\partial}{\partial r}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial r}\right\}+\frac{\partial}{\partial z}\left\{\frac{1}{r\mu}\frac{\partial\varphi}{\partial z}\right\}+\frac{\omega^2\epsilon}{c^2r}\cdot\varphi=0. \tag{3.05}$$

Suppose that

$$\mu=\mu_1(r)\cdot\mu_2(z), \tag{3.06}$$

$$\epsilon=\epsilon_1(r)\cdot\epsilon_2(z), \tag{3.07}$$

$$\varphi=R(r)\cdot Z(z), \tag{3.08}$$

then

$$0=\frac{r\mu_1(r)}{R}\frac{d}{dr}\left\{\frac{1}{r\mu_1(r)}\frac{dR}{dr}\right\}+\frac{\mu_2(z)}{Z}\frac{d}{dz}\left\{\frac{1}{\mu_2(z)}\frac{dZ}{dz}\right\} +\frac{\omega^2}{c^2}\mu_1(r)\epsilon_1(r)\mu_2(z)\epsilon_2(z). \tag{3.09}$$

This equation is separable if either

$$\mu_1(r)\cdot\epsilon_1(r)=k_r \ (k_r\text{-separability}), \tag{3.10}$$

$$\mu_2(z)\cdot\epsilon_2(z)=k_z \ (k_z\text{-separability}). \tag{3.11}$$

For k_r-separability

$$\frac{r\mu_1(r)}{R}\frac{d}{dr}\left\{\frac{1}{r\mu_1(r)}\frac{dR}{dr}\right\}+\lambda_1^2=0, \tag{3.12}$$

$$0=\frac{\mu_2(z)}{Z}\frac{d}{dz}\left\{\frac{1}{\mu_2(z)}\frac{dZ}{dz}\right\}-\left\{\lambda_1^2-\frac{\omega^2}{c^2}k_r\mu_2(z)\epsilon_2(z)\right\}. \tag{3.13}$$

While for k_z-separability

$$\frac{r\mu_1(r)}{R}\frac{d}{dr}\left\{\frac{1}{r\mu_1(r)}\frac{dR}{dr}\right\}+\frac{\omega^2}{c^2}k_z\mu_1(r)\epsilon_1(r)=\lambda_2^2, \tag{3.14}$$

$$\frac{\mu_2(z)}{Z}\frac{d}{dz}\left\{\frac{1}{\mu_2(z)}\frac{dZ}{dz}\right\}+\lambda_2^2=0. \tag{3.15}$$

Up to this point no use has been made of Eq. (3.01) and (3.02), and the equations might be solved by accurate or approximate values of the functions $\mu_1(r)$, $\epsilon_1(r)$ in Eq. (3.14) or arbitrary values of $\mu_2(z)$, $\epsilon_2(z)$ in (3.13). If μ and ϵ have the values given in (3.01) and (3.02), $\mu\epsilon$ is a

constant, and the equations for the factors μ_1, μ_2, ϵ_1, and ϵ_2, become

$$\mu_1 = Mr^{2k-2}, \qquad \mu_2 = e^{2az}, \tag{3.16}$$

$$\epsilon_1 = \mu_0\epsilon_0/Mr^{2k-2}, \quad \epsilon_2 = e^{-2az}. \tag{3.17}$$

Both separability conditions are met, so either (3.12), (3.13) or (3.14), (3.15) can be used. Substituting these values in (3.12) and (3.13) gives

$$\frac{Mr^{2k-1}}{R}\frac{d}{dr}\left\{\frac{1}{Mr^{2k-1}}\frac{dR}{dr}\right\}+\lambda_1^2=0 \tag{3.18}$$

and

$$\frac{e^{2az}}{Z}\frac{d}{dz}\left\{e^{-2az}\frac{dZ}{dz}\right\}-\{\lambda_1^2-\omega^2k_r/c^2\}=0$$

or

$$0=Z''-2aZ'-\{\lambda_1^2-\omega^2\mu_0\epsilon_0/c^2\}Z. \tag{3.19}$$

This is immediately solved by

$$Z=\exp pz \tag{3.20}$$

where

$$p=a\pm i(\omega^2\mu_0\epsilon_0/c^2-\lambda_1^2-a^2)^{\frac{1}{2}}. \tag{3.21}$$

Writing $R=r^kS(r)$ Eq. 3.18 becomes

$$S''+1/r\cdot S'+\{(\lambda_1)^2-k^2/r^2\}S=0. \tag{3.22}$$

Thus it is clear that the radial function is

$$R=r^k\{AJ_k(\lambda_1 r)+BN_k(\lambda_1 r)\} \tag{3.23}$$

and

$$\varphi=r^k\{AJ_k(\lambda_1 r)+BN_k(\lambda_1 r)\}e^{pz}. \tag{3.24}$$

In this problem we started with the arbitrary distribution of μ and ϵ given by (3.01) and (3.02), and showed the solution could be expressed in terms of Bessel Functions and exponentials, a generalization of the solution applying when μ and ϵ are constant, which is

$$\varphi_0=r\{AJ_1(\lambda_1 r)+BN_1(\lambda_1 r)\}e^{pz}. \tag{3.25}$$

The non-vanishing field components are

$$E_\theta=-i\omega/c\cdot r^{k-1}\{AJ_k(\lambda_1 r)+BN_k(\lambda_1 r)\}e^{pz}, \tag{3.26}$$

$$B_r=-pr^{k-1}\{AJ_k(\lambda_1(r))+BN_k(\lambda_1 r)\}e^{pz}, \tag{3.27}$$

$$B_z=\lambda_1 r^{k-1}\{AJ_{k-1}(\lambda_1 r)+BN_{k-1}(\lambda_1 r)\}e^{pz}, \tag{3.28}$$

since Bessel functions obey

$$[dZ_k/d(\lambda_1 r)]=Z_{k-1}-k/\lambda_1\cdot Z_k r^{-1}.$$

The solutions agree with the classic solutions for a region where μ and ϵ are constant, as is seen on taking $k=1$, $a=0$.

This choice of the function $\mu(r, z)$ leads to an infinite value of either μ or ϵ on the axis. For this reason it is necessary to specify that the medium does not extend all the way in to the axis. An interesting case is that of a co-axial line where the space between the conductors is filled with a medium specified by Eqs. (3.01) and (3.02). If the medium fills a co-axial line, and is bounded by highly conducting metal cylinders of radii r_0 and r_1, then the condition for the vanishing of the tangential electric field is

$$J_k(\lambda_1 r_0)/N_k(\lambda_1 r_0)=J_k(\lambda_1 r_0)/N_k(\lambda_1 r_0). \tag{3.29}$$

The roots in λ_1 of this equation determine the values of the propagation constant p. This constant is complex at all angular frequencies above

$$f_c=1/2\pi\cdot c\{(\lambda_1^2+a^2)/\mu_0\epsilon_0\}^{\frac{1}{2}}.$$

Below this frequency the propagation constant is a real number.

Section 4. TM Waves of Type C

Since to each solution of the field equations representing a TE wave $\varphi_3=\varphi(\mu, \epsilon)$ there is a corresponding TM wave generated by $\psi_3=\psi(\epsilon, \mu)$, it is possible to write down immediately the TM solutions corresponding to (3.01), (3.02), and (3.24), namely

$$\epsilon=Mr^{2k-2}e^{2az}, \tag{4.01}$$

$$\mu=\mu_0\epsilon_0/\epsilon, \tag{4.02}$$

$$\psi=r^k\{CJ_k(hr)+DN_k(hr)\}e^{pz}, \tag{4.03}$$

where

$$p=a\pm i\{\omega^2\mu_0\epsilon_0/c^2-h^2-a^2\}^{\frac{1}{2}}. \tag{4.04}$$

The non-vanishing field components are

$$H_\theta=-\omega/c\cdot r^{k-1}\{CJ_k(hr)+DN_k(hr)\}e^{pz}, \tag{4.05}$$

$$D_r=-ipr^{k-1}\{CJ_k(hr)+DN_k(hr)\}e^{pz}, \tag{4.06}$$

$$D_z=ih\cdot r^{k-1}\{CJ_{k-1}(hr)+DN_{k-1}(hr)\}e^{pz}. \tag{4.07}$$

If the medium described by Eqs. (4.01) and (4.02) fills the space between two co-axial conductors of radii r_0 and r_1, the boundary conditions $D_z(r_0)=D_z(r_1)=0$ give

$$J_{k-1}(hr_0)/N_{k-1}(hr_0)=J_{k-1}(hr_1)/N_{k-1}(hr_1). \tag{4.08}$$

These results obviously reduce to the classic results for a co-axial guide when $k=1$, $a=0$.

Section 5. TEM Waves

For the case of a co-axial line filled with a medium described by (4.01) and (4.02), TEM waves can be propagated if $D_z=0$. The fields for this case are described by

$$\partial/\partial z\{1/\epsilon\cdot\partial\psi/\partial z\}+\omega^2/c^2\cdot\mu\psi=0, \tag{2.02}$$

$$D_r=-i/r\cdot\partial\psi/\partial z, \quad H_\theta=-\omega/rc\cdot\psi. \tag{2.03}$$

For this particular problem

$$\frac{d}{dz}\left\{e^{-2az}\frac{d\psi}{dz}\right\}+\frac{\omega^2}{c^2}\cdot\mu_0\epsilon_0 e^{-2az}\psi=0. \tag{5.01}$$

If

$$\psi=e^{pz},$$

$$p=a\pm i\{\omega^2\mu_0\epsilon_0/c^2-a^2\}^{\frac{1}{2}}. \tag{5.02}$$

When $\omega^2<a^2c^2/\mu_0\epsilon_0$, the propagation constant p is real. If ω is very small $p=\frac{1}{2}\omega^2\mu_0\epsilon_0/ac^2$ for waves propagated

in the positive z direction. The impedance ratio becomes, from (2.04)

$$Z=iac/\omega\epsilon\mp\{\mu_0\epsilon_0/\epsilon^2-a^2c^2/\omega^2\epsilon^2\}^{\frac{1}{2}}. \quad (5.03)$$

The $\pm$ sign indicates that the impedance ratio is different for right and left traveling waves, in magnitude as well as in sign. The lower sign refers to waves traveling in the positive z direction. (The time factor is $e^{+j\omega t}$, since $\partial_4=\omega/c$.)

For waves traveling in the positive direction

$$D_r=pAe^{pz}/ir, \quad H_\theta=-\omega Ae^{pz}/cr. \quad (5.04)$$

For a co-axial line, the characteristic impedance defined as

$$Z_0=V/I=\int_{r_0}^{r_1} E_r\cdot dr \Big/ \int_0^{2\pi} H\cdot ds$$

is of greater importance than the impedance ratio $Z=E_r/H_\theta$. In this case

$$Z_0=\frac{1}{2\pi M}\left\{\left[\mu_0\epsilon_0-\frac{a^2}{\omega^2c^2}\right]^{\frac{1}{2}}+\frac{ia}{\omega c}\right\} \times\left\{\frac{{r_1}^{2-2k}-{r_0}^{2-2k}}{2-2k}\right\}e^{-2az}. \quad (5.05)$$

When $k=1$, $a=0$, this reduces to

$$Z_0(1,0)=1/2\pi\cdot(\mu_0/\epsilon_0)^{\frac{1}{2}}\log_e r_1/r_0 \quad (5.06)$$

in agreement with the classic impedance of a co-axial line.

Section 6. Waves in a Region Where ϵ Varies but μ is Constant

The previous examples have dealt with cases where both μ and ϵ were functions of position. In the Type C examples, the refractive index, $(\mu\epsilon)^{\frac{1}{2}}$, was constant, although μ and ϵ varied. Another type of example, of evident practical importance, is the case where μ is constant, but ϵ is some function of position. Referring to Table III, when $dx^1=dz$, $dx^2=dr$, $dx^3=d\theta$, the propagation equation was

$$\frac{\partial}{\partial z}\left\{\frac{1}{r\epsilon}\frac{\partial\psi}{\partial z}\right\}+\frac{\partial}{\partial r}\left\{\frac{1}{r\epsilon}\frac{\partial\psi}{\partial r}\right\}+\frac{\omega^2}{c^2}\frac{\mu}{r}\psi=0. \quad (6.01)$$

In this case μ is constant, but ϵ may be taken to be the product of two factors

$$\epsilon=\epsilon_1(r)\cdot\epsilon_2(z). \quad (6.02)$$

Factorizing ψ into $\psi=RZ$, (6.01) becomes

$$\frac{r\epsilon_1}{R}\frac{d}{dr}\left\{\frac{1}{r\epsilon_1}\frac{dR}{dr}\right\}+\frac{\epsilon_2}{Z}\frac{d}{dz}\left\{\frac{1}{\epsilon_2}\frac{dZ}{dz}\right\}+\frac{\omega^2\mu\epsilon_1\epsilon_2}{c^2}=0. \quad (6.03)$$

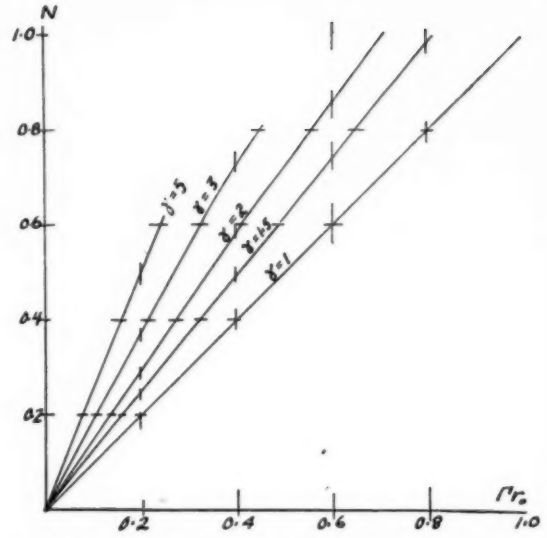


FIG. 5. Relation between the propagation constant Γ and the quantity $N=\omega r_0(\mu\epsilon_0)^{\frac{1}{2}}/c$ for a co-axial line where the ratio of outer to inner diameter is γ, filled with a medium where $\mu=$constant, $\epsilon=\epsilon_0(r_0/r)^2$.

Applying the operator $\partial^2/\partial r\partial z$ to (6.03) gives

$$0=d\epsilon_1/dr\cdot d\epsilon_2/dz.$$

The equation is separable if ϵ is a function of r only or of z only. For co-axial line propagation, $\epsilon=\epsilon(z)$ leads to the well known results for a tapered transmission line.[5] Therefore only the case $\epsilon=\epsilon(r)$ will be considered. In this case, with $\epsilon_2=1$

$$\frac{r\epsilon(r)}{R}\frac{d}{dr}\left\{\frac{1}{r\epsilon}\frac{dR}{dr}\right\}+\left(\frac{\omega^2\mu\epsilon(r)}{c^2}+p^2\right)=0, \quad (6.04)$$

$$Z''=p^2Z. \quad (6.05)$$

The Z factor is clearly

$$Z=e^{\pm pz}. \quad (6.06)$$

The R factor depends on the chosen distribution of ϵ. For it the equation may be written

$$R''-R'\cdot d/dr\log r\epsilon+(p^2+\omega^2\mu\epsilon(r)/c^2)R=0. \quad (6.07)$$

In the particular case where $\epsilon=\epsilon_0(r_0/r)^{2}$** is chosen, the Eq. (6.07) becomes

$$R''+1/r\cdot R'+(p^2+\omega^2\mu\epsilon_0{r_0}^2/c^2r^2)R=0 \quad (6.08)$$

whose solution is

$$R=AJ_{i\omega r_0(\mu\epsilon_0)^{\frac{1}{2}}/c}(pr)+BN_{i\omega r_0(\mu\epsilon_0)^{\frac{1}{2}}/c}(pr). \quad (6.09)$$

This leads to a surprising exchange of role between p and $\omega^2\mu\epsilon/c^2$.

[5] Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc. New York, 1943), p. 205.

** Again, a more general problem could be solved by taking $\omega^2\mu\epsilon/c^2=A+Br^{-2}$.

To derive an approximate expression for the relation between $\Gamma = ip$ and $N = \omega r_0/c.\ (\mu\epsilon_0)^{\frac{1}{2}}$ for low frequencies or small radii, it is necessary to use the usual positive power expansion for Bessel's functions. Since the solution contains Bessel's functions of non-integral order (order being imaginary), the solutions of the propagation equation may be rewritten as

$$\psi = \{AJ_{iN}(pr) + BJ_{-iN}(pr)\}e^{pz} \tag{6.10}$$

where

$$J_{iN}(pr) = \{\sum_0^\infty a_{2m}\cdot(pr)^{iN+2m}\}/2^{\underline{iN}}\underline{|iN}, \tag{6.11}$$

$$J_{-iN}(pr) = \{\sum_0^\infty a_{2m}^*\cdot(pr)^{-iN+2m}\}/\underline{2^{-iN}|} - iN \tag{6.12}$$

and

$$a_{2m} = \underline{|Ni}/(\underline{|m|}\,\underline{m+iN}(2i)^{2m}). \tag{6.13}$$

At the surface of a cylindrical conductor of radius r_0 the boundary condition is

$$-B/A = J_{iN}'(pr_0)/J_{-iN}'(pr_0). \tag{6.14}$$

For a co-axial line with conductors of radii r_0 and $r_1 = \gamma r_0$, calling $u = pr_0$

$$J_{iN}'(u_0)/J_{-iN}'(u_0) = J_{iN}'(\gamma u_0)/J_{-iN}'(\gamma u_0). \tag{6.15}$$

Using the expansions (6.11) and (6.12) and the values of a_{2m} from (6.13) in (6.15) finally gives the boundary conditions of the co-axial line in the form

$$\sum_{s=0}^\infty u_0^{2s}A_s(N) = 0 \tag{6.16}$$

with

$$A_s(N) = -i\sum_{m=0}^{s}\{a_{2m}a_{2s-2m}^*(2m+iN) \times(2s-2m-iN)\gamma^{2s-2m-iN} - a_{2m}^*a_{2s-2m}(2m-iN)(2s-2m+iN)\gamma^{2s-2m+iN}. \tag{6.17}$$

Introducing the expression

$$\rho_{m,s}e^{i\varphi_{m,s}} = \underline{|iN|}\,\underline{-iN}(2m+iN) \times(2s-2m-iN)/\underline{|m+iN|}\,\underline{s-m-iN} \tag{6.18}$$

then by (6.17), (6.13), and (6.18)

$$iA_s = (-)^s 2^{1-2s}\sum_{m=0}^{s}\gamma^{2s-2m}\rho_{m,s} \times\sin(\varphi_{m,s}-\sigma N)/\underline{|m|}\,\underline{s-m} \tag{6.19}$$

with

$$\sigma = \log_e\gamma.$$

For small radii or low frequencies, we have

$$0 = A_0(N) + u_0^2A_1(N) \tag{6.20}$$

leading to the approximate equation

$$-u_0^2 = \Gamma^2r_0^2 = 4N(1+N^2)/\{(\gamma^2+1)N + (\gamma^2-1)(2+N^2)\cot\sigma N\} \tag{6.21}$$

and for extremely small values of N

$$\Gamma r_0 \simeq \{2\sigma/(\gamma^2-1)\}^{\frac{1}{2}}\cdot N.$$

It is interesting to note that Γr_0 tends to the usual value for a co-axial line as $\gamma \to 1$, for evidently, with $\sigma = \log\gamma$,

$$\lim\gamma\to 1\ \Gamma^2r_0^2 = \omega^2\mu\epsilon_0r_0^2/c^2.$$

Figure 5 shows the relation between N and Γr_0 for low frequencies.

It is hoped that these examples will serve to illustrate the application of this theory to radio and communication problems.

The author wishes to thank Hilda E. Smith for assistance with the manuscript and proof.

Hypersonic Research Facilities at the Ames Aeronautical Laboratory

VICTOR I. STEVENS
National Advisory Committee for Aeronautics, Ames Aeronautical Laboratory, Moffett Field, California
(Received June 14, 1950)

Two wind tunnels recently completed at the Ames Laboratory of the NACA—the 10- by 14-in. supersonic wind tunnel and the supersonic free-flight wind tunnel—are capable of providing test Mach numbers well in excess of 5. Since at this time such facilities are relatively uncommon, this paper is primarily devoted to a description of these wind tunnels and their associated equipment.

The treatment of the 10- by 14-in. supersonic wind tunnel is concerned for the most part with the design and operation of the nozzle and diffuser which provides Mach numbers of 3.5 to about 8. Some mention is also made of the pressure measuring and flow visualization equipment. In the case of the supersonic free-flight wind tunnel high test Mach numbers are obtained by firing models at high speed into an oncoming supersonic air stream. Consequently the main emphasis is placed upon the launching techniques, the available range of test conditions, and the methods of obtaining data.

INTRODUCTION

ONLY a few years ago, interest in high Mach number flows was largely academic and as a result relatively little effort was directed toward solving the problems associated with generating and studying such flows. Recently, however, interest in these flows has greatly intensified and much more pressure has been brought to bear to develop research equipment which will facilitate the study of air flow at Mach numbers of 5 and above. The NACA at its Langley, Lewis, and Ames Laboratories and other research groups have been devoting increasing effort to the development of such equipment.

Perhaps the greatest obstacles to the progress of hypersonic research are the difficulties encountered in providing the high compression ratios required for the generation of high Mach number flows and in providing the instrumentation to cover the broad ranges of temperature, pressure, density, etc., that must be observed and recorded in the investigation of hypersonic flow phenomena. Two wind tunnels recently completed at the Ames Laboratory—the 10- by 14-in. supersonic wind tunnel and the supersonic free-flight wind tunnel—incorporate techniques which at least partially solve some of these problems. It is the purpose of this paper to give a brief description of these wind tunnels and their associated equipment.

10- BY 14-IN. SUPERSONIC WIND TUNNEL

The 10- by 14-in. supersonic wind tunnel has only recently been completed and is now being calibrated. It is a closed-throat, non-return, continuous-flow tunnel capable of producing any air-stream Mach number from 3.5 to something in excess of 8. The wind tunnel was designed to utilize existing compressors having a maximum exhaust pressure of only 6 atmos. absolute; consequently, it was important to keep the compression ratio required for a given Mach number as low as possible. Use of a conventional laval nozzle followed by a subsonic divergent-wall diffuser as shown in the upper half of Fig. 1 would require compression ratios of the order of 200 for a Mach number of 8. With a supply pressure of only 6 atmos. the exit pressure would have to be less than 0.03 atmos.—a value which is unacceptably low, for it would require very bulky vacuum equipment. Most of this compression ratio is required to overcome the head loss through the normal shock; therefore, use of a supersonic convergent-divergent diffuser (see lower half of Fig. 1) which decelerates the air before the normal shock, appeared advantageous to reduce this head loss and thereby reduce the compression ratio required and increase the allowable exit pressure. The limited data available at the time this design was undertaken indicated that small gains could be realized through use of this type diffuser for Mach numbers of 2 to 3, but there was little or no information available to direct the design of such a diffuser

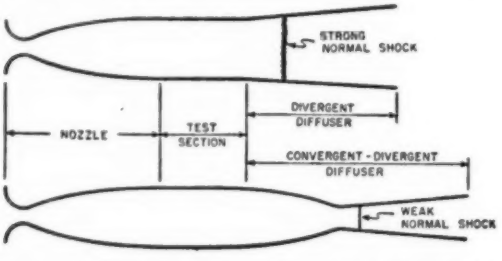


FIG. 1. Diagram of supersonic nozzles equipped with two types of diffusers.

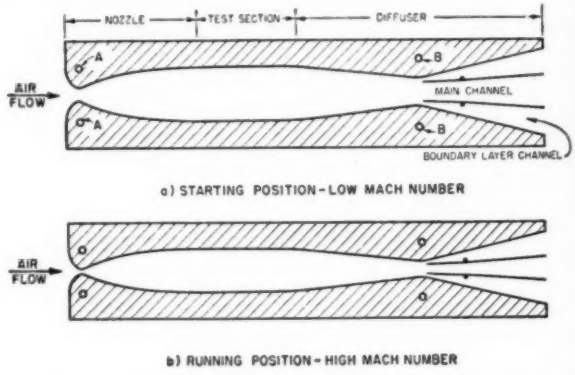


FIG. 2. Arrangement of nozzle, test section, and diffuser for the Ames 10- by 14-in. supersonic wind tunnel.

for Mach numbers as high as 8. Therefore, in 1947 an investigation of a 1- by 1-in. pilot tunnel was undertaken to develop a nozzle and diffuser for the 10- by 14-in. supersonic tunnel.

Nozzle and Diffuser

The nozzle and diffuser configuration developed is shown schematically in Fig. 2. The side walls are flat and parallel and all contour variation is in the top and bottom blocks. These blocks are rigid and are supported on pins at A and B. Jacks attached to these pins control the vertical position of the blocks and thus control the nozzle and diffuser throat openings. In a somewhat oversimplified view, opening or closing the jacks at A lowers or raises the test-section Mach number, whereas the position of jacks at B determines the location and strength of the normal shock.

Just downstream of the diffuser throat are located boundary-layer scoops. The investigations in the pilot tunnel revealed that without these scoops the normal shock boundary-layer interaction was very severe and the resulting erratic behavior of the normal shock and boundary layer made it necessary to force the normal shock far downstream of the diffuser throat in order to preserve supersonic flow ahead of the diffuser throat. The extra compression ratio required to keep the normal shock a safe distance downstream nullified, to a large extent, the gains achieved by a convergent-divergent diffuser. Removal of the boundary layer, however, stabilized the normal shock position so that supersonic flow ahead of the diffuser throat could be maintained with the shock only a short distance downstream of the scoop opening. Boundary-layer removal also greatly reduced the separation behind the normal shock and thus improved the efficiency of the subsonic diffuser.

The operation of this nozzle can probably be better understood by following through the sequence of starting supersonic flow and then changing the Mach number. The starting position is as shown in the top of Fig. 2. With the blocks open to this position, as the pressure builds up, supersonic flow at a Mach number of approximately 3.5 is established in the test section and the normal shock is pushed out of the test section through the converging section of the diffuser and into the divergent section. To establish flow with this nozzle configuration requires a compression ratio of about 6 which, as would be expected, is about the same as that required to obtain a Mach number of 3.5 in a wind tunnel with a conventional divergent diffuser. It can be shown[1] that supersonic flow in the test section will not start regardless of compression ratio available, if the second throat is not open beyond a certain minimum. For this wind tunnel the second throat is opened so that when supersonic flow is established the Mach number at the second throat is about 3. However, after flow is established the diffuser throat may be closed to give a throat Mach number of approximately 2.5. To increase the test-section Mach number, the blocks are closed as shown in the lower half of Fig. 2. With the nozzle throat closed to an opening of about 0.05 in. the test-section Mach number is 8 and the Mach number at the diffuser throat is about 3.5. The compression ratio required for this Mach number is about 30:1 which is considerably less than the estimated 200:1 required for a conventional divergent diffuser.

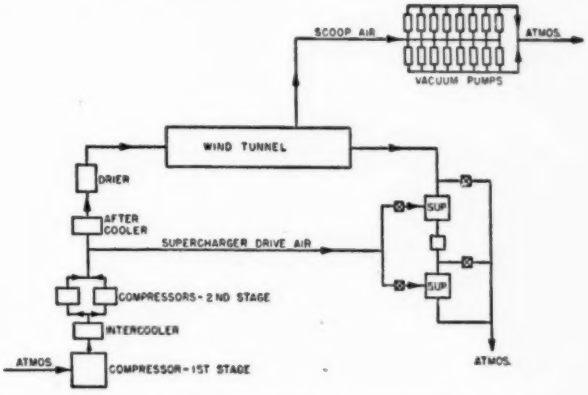


FIG. 3. Air-flow diagram for the Ames 10- by 14-in. supersonic wind tunnel.

Air Supply

The high pressure air for this wind tunnel is supplied by the equipment shown on the left side of Fig. 3. This equipment consisting of centrifugal compressors, coolers, and a silica-gel drier provides air at the nozzle entrance at a maximum total pressure of 6 atmos., a temperature of approximately 100°F and an absolute humidity of 0.0001 lb. water/lb. air. The air after passing through the nozzle and test section exhausts into either the boundary-layer scoops or the main diffuser. The low energy boundary-layer air is drawn off through the scoops by a battery of rotary vacuum pumps and discharged to the atmosphere. The air exhausting through the main diffuser is discharged directly to the atmosphere at test Mach numbers below 4.5. At Mach numbers above 4.5 this air is drawn through either one or two stages of turbine-driven superchargers to provide the required compression ratios. Since the mass flow through the tunnel decreases with increasing Mach number, it is convenient to utilize the excess air from the primary compressors to drive the turbines on the superchargers. With both superchargers operating the total compression ratio across the main passage is 55 which is more than adequate for test Mach numbers of 8.

Instrumentation

The model support and force-measuring equipment for this tunnel are very similar to that used in other supersonic wind tunnels. The models are supported

[1] A. Kantrowitz and C. du P. Donaldson, Preliminary Investigation of Supersonic Diffusers, NACA ACR L5D20 (1945).

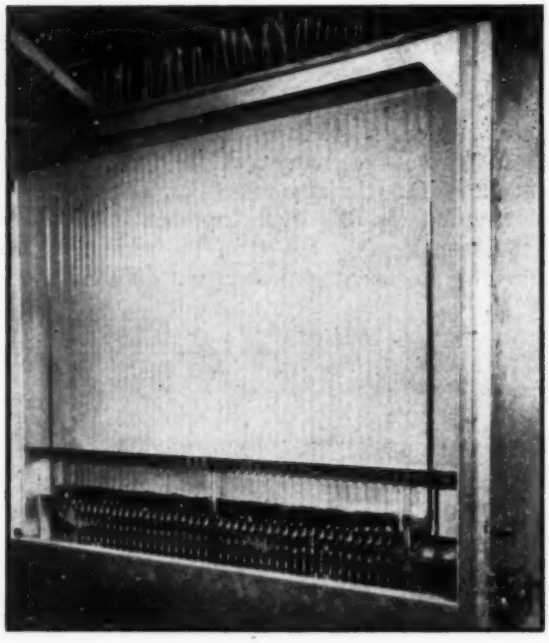


FIG. 4. Manometer board incorporating McLeod-type gauges for measuring low pressures.

from the rear on a sting which is in turn supported by a strut passing through the sidewalls. The forces and moments acting on the model are measured using a conventional strain-gauge balance system.

In contrast, the low pressures and densities existing in the wind-tunnel test section at high Mach numbers required deviations in design from the conventional pressure-measuring and flow-visualization equipment. For a given stagnation pressure the test-section static and dynamic pressures fall very rapidly with Mach number so that at a test Mach number of 8 and a stagnation pressure of 6 atmos. the static and dynamic pressures are 0.000612 and 0.0275 atmos., respectively (or 0.018 and 0.83 in. Hg). To measure pressures of this magnitude the conventional U-tube-type manometer board has been replaced by a board incorporating 40 McLeod-type gauges. This manometer board is shown in Fig. 4. Up to 40 pressures can be simultaneously recorded by photographing this board. Using these photographic records absolute pressures of 0.001 atmos. can be read to approximately one percent accuracy.

Flow about models is observed using a center-of-curvature schlieren apparatus which has much greater sensitivity to density gradients than the more common schlieren system. With the center-of-curvature schlieren the effective source and knife are located on one side of the test section at the center of curvature of the principal mirror. The mirror is located immediately adjacent to the test section on the side opposite the source and knife. Thus the light passes through the test section twice and there is no collimated light beam as in the common schlieren. The principal mirror for the 10- by 14-in. supersonic tunnel has a diameter of 1½ ft. and a radius of curvature of 50 ft. Although the schlieren system incorporating this mirror has not yet been fully checked, it is expected that the system will have adequate sensitivity over the Mach number range of this wind tunnel.

THE SUPERSONIC FREE-FLIGHT WIND TUNNEL

The supersonic free-flight wind-tunnel structure was completed early in 1949; however, the installation of the equipment required to obtain aerodynamic data was not completed until late in the fall of the same year. The major calibrations are complete and the facility is

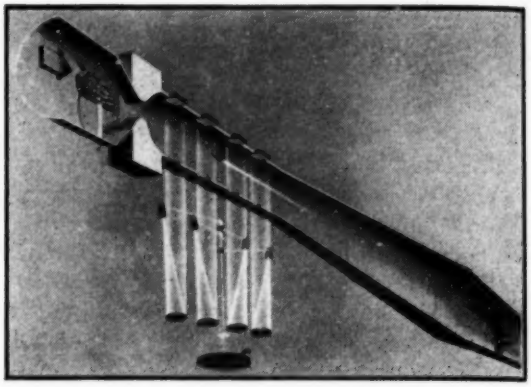


FIG. 5. Arrangement of the Ames supersonic free-flight wind tunnel.

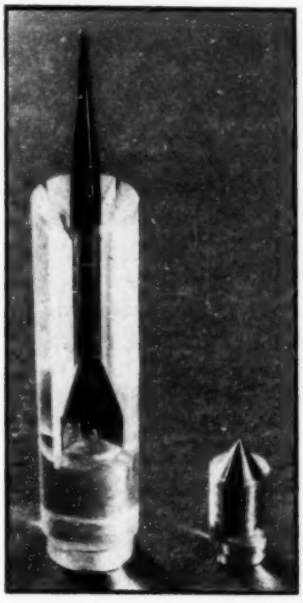
now being used to study the characteristics of missile-type configurations. Although the maximum test Mach number obtained to date is only slightly in excess of 8, the facility is capable of providing data for Mach numbers ranging from less than 1 to approximately 15. The necessity for high compression ratios at high test Mach numbers is obviated by generating an air-stream Mach number of only 2 or 3 in the wind-tunnel test section and providing the remainder of the total Mach number by launching the model at high velocities into the oncoming air stream. Since the air in passing through the nozzle is expanded only to a Mach number of 2 or 3, the air density and consequently the Reynolds number for a given model size are relatively high.

Data that can be obtained in this wind tunnel include the model static force characteristics which are obtained in conventional wind tunnels, as for example, lift, drag, and moment. However, the techniques for measuring these characteristics parallel those used in aeroballistic ranges.[2] The dynamic behavior of the models can also be obtained directly in this wind tunnel. The methods for obtaining these data can probably be better understood after examining the wind tunnel and related equipment.

[2] A. C. Charters, "The spark photography range," a paper given at the International Congress for Applied Mechanics (September, 1946).

Wind Tunnel and Launching Guns

The principal components of the wind tunnel proper are shown in Fig. 5. The wind tunnel is an intermittent-operation blow-down type with the air being supplied from the 12-ft. pressure wind tunnel at a maximum pressure of 6 atmos. The air flows from left to right as shown in Fig. 5 and passes, in turn, through the settling chamber, the nozzle, the test section, and the diffuser. After diffusion the air is ejected directly to the atmosphere. The test section is 1 ft. wide, 2 ft. deep, and 18 ft. long. The Mach number of the air flow in the test section is, of course, a function of the shape of the nozzle blocks. At present, the wind tunnel is equipped with two sets



(a) (b)

FIG. 6. Models and sabots as launched from guns. (a) Fin-stabilized. (b) Spin-stabilized.

of interchangeable blocks which give an air-stream Mach number of either 2 or 3. The models are fired from guns located in the diffuser and after passing through the test section and nozzle are stopped by the model catcher in the reservoir. The model catcher is a steel cylinder $2\frac{1}{2}$ ft. in diameter and about 6 ft. long which is tightly packed with cotton waste backed up by wood and steel.

Launching of models at high velocity from guns requires techniques which in the past have been quite foreign to wind tunnels. The techniques which we use at present are closely patterned after those developed at the Naval Ordnance Laboratory, the Watertown Arsenal, and the Ballistic Research Laboratories at Aberdeen Proving Ground. While in the gun the models are supported in carriers or sabots illustrated in Fig. 6. The purpose of the sabot is to provide a gas seal and at the same time protect the model from the hot powder gases and to keep it properly aligned during launching. There are two basic types of sabots which we are using at present—one for fin-stabilized models

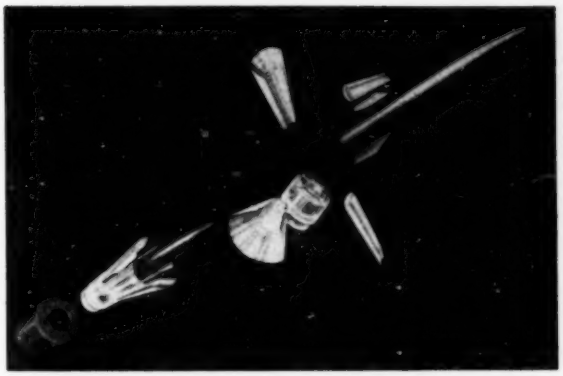


FIG. 7. Drawing showing sabot in two stages of separation from fin-stabilized model.

and one for spin-stabilized models (see Fig. 6). Both types are so constructured that when the sabot and model emerge from the gun, the sabots separate from the model and leave it free to fly on undisturbed through the wind tunnel. Sabot separation on a fin-stabilized model is illustrated in Fig. 7.

The launching guns available for use in this wind tunnel range in size from caliber .22 through .50, 20 mm, 37 mm, and 3 in. The launching velocities are varied by changing the powder charges loaded into the cartridge case—the limiting velocity being determined by the maximum pressure that the gun can withstand. In the .220 Swift, which is a sporting rifle, we have realized launching velocities of 5000 ft./sec. which, with the Mach number 2 nozzle blocks installed, gives a test Mach number of 8. It may not at first be evident how a launching velocity of 5000 ft./sec., which is a nominal sea-level Mach number of $4\frac{1}{2}$, can add to the test section of 2 and yield 8. This, however, is traceable to the fact that the air in expanding through the nozzle is chilled to temperatures of about $-150°$F. At these low temperatures the speed of sound is of course lowered; hence, the launching velocity of 5000 ft./sec. in itself produces a Mach number of 6 which when added to the air stream Mach number of 2 gives 8. Special test barrels are being procured which should permit launching velocities of the order of 8000 ft./sec. With the Mach number 3 nozzle blocks installed in the wind tunnel this launching velocity would give a test Mach number of 15. It should be noted that a compression ratio of only 4 is required for tests at this Mach number and that to realize the same Mach number and air density in the region of the model in a conventional wind tunnel would require a stagnation pressure of more than 100,000 atmos.!

Due to the comparatively high air density in the wind-tunnel test section the test Reynolds numbers are very high for a given model size. For example, if a 6-in. model were launched in this tunnel to give a test Mach number of 7, the resulting Reynolds number would correspond to that obtained on a 50-ft. missile flying at the same Mach number at an altitude of

100,000 ft. Such a rocket is comparable in size to a V-2. In contrast, to obtain the same Reynolds number in a conventional tunnel operating at the same stagnation pressure (6 atmos.) would require a model over 10 ft. long.

Equipment and Techniques for Measuring Model Characteristics

As stated previously, the aerodynamic characteristics of the models are determined in this wind tunnel using techniques which parallel those used in aeroballistic

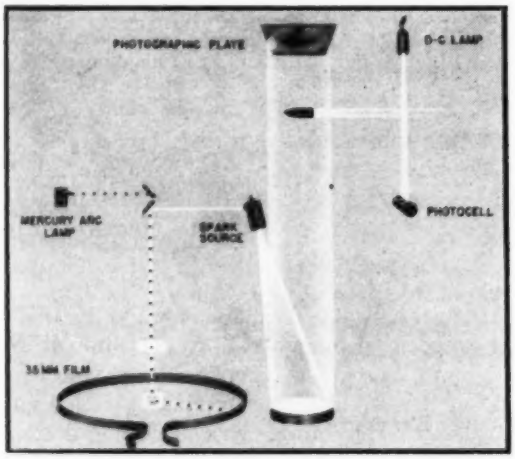


FIG. 8. Diagram illustrating operation of time-distance equipment.

firing ranges. To obtain drag, for example, the time-distance history of the model is recorded as the model flies through the test section and from this history the velocity, deceleration, and hence drag can be evaluated. In this wind tunnel the time-distance history is obtained using a four-station shadowgraph and a chronograph as shown in Fig. 5. As the model flies through the wind tunnel, each station in turn takes a shadowgraph of the model and a distance scale so that the distance the model travels from station to station can be measured on the shadowgraph. At the same instant each shadowgraph is taken, the chronograph records the time.

The operation of this equipment is more easily followed using Fig. 8, which isolates one of the shadowgraph stations and the principal components of the chronograph. When the model leaves the launching gun and enters the test section, it first encounters a light beam which projects through the wind tunnel and impinges on a photoelectric cell. As the model interrupts part of this light beam the output of the photoelectric cell is modified. The resulting signal is amplified and passed through a suitable time delay to allow the model to travel to the shadowgraph station. When this time delay has elapsed, the spark source is triggered and the resulting light forms a shadow image of the model on the photographic plate. Although not shown in Fig. 8, the distance scale is suspended below and along one edge of the photographic plate so that its image is also formed on the plate. When the spark is triggered, part of the light escapes from the side of the spark housing where it is picked up by a mirror and reflected along a vertical axis to a rotating prism which in turn reflects the light out to the 35-mm film on the periphery of the chronograph. Thus a latent image is formed on the chronograph film at the instant the shadowgraph is taken. As the model reaches the second, third, and fourth stations this same procedure is repeated; however, since the prism in the chronograph is rotating, the spark images are spaced approximately 90° apart on the 35-mm film. To provide a time base an H-6 mercury arc lamp is pulsed at frequencies which are regulated at any desired frequency from 10,000 to 100,000 cycles/sec. and the resulting pulsed light superimposed on the chronograph film. Since the pulse rate can be varied from 10,000 to 100,000 cycles/sec. the resulting interval between pips on the film varies from 100 to 10 μsec.

A sample of the shadowgraph and chronograph records is shown on Fig. 9. The shadowgraph was obtained on a 60° cone-cylinder model at a test Mach number of 7. The sharply defined notches of the distance scale can be seen on the upper side of the shadowgraph. At the bottom of Fig. 9 is a short segment of the chronograph record. The station marker is at the center of the section and the evenly spaced pips are the time base—each interval representing 20 μsec.

The accuracy requirements on measuring distance and time are rather stringent. For example, to restrict

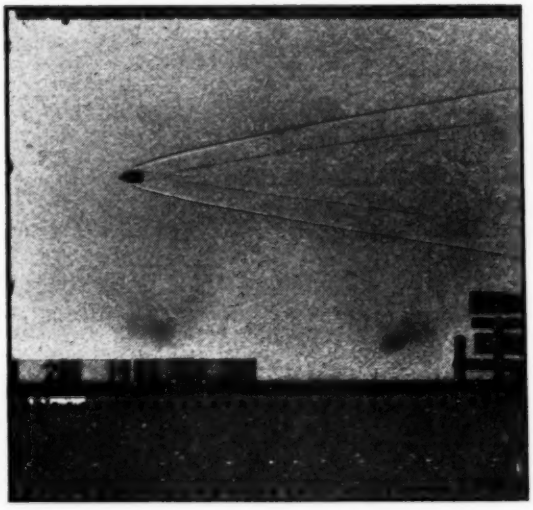


FIG. 9. Sample shadowgraph and chronograph records.

errors in drag to one percent requires in some cases distance measurements to the accuracy of 0.001 in. and time to 0.01 μsec. The requirement that time be measured to this order of accuracy becomes more significant when it is realized that in 0.01 μsec. light travels a dis-

tance of about 10 ft. It is evident, then, that in the design of equipment of this type the speed of light and light-path lengths must be considered.

Several other forms of aerodynamic data can be obtained using the time-distance record. With three lateral shadowgraph stations (omitted from Fig. 5 for clarity) and the four vertical stations already described, it is possible to define the three-dimensional flight path of the model. From this flight path, lateral and vertical accelerations, hence side force and lift, can be obtained. It is also possible to obtain base pressures using the method of characteristics to determine the conditions just ahead of the model base and using any one of the shadowgraphs to obtain the flow expansion around the base.[3] Stability parameters such as lift-curve slope, aerodynamic-center position, damping-in-roll, etc., can be evaluated by disturbing the model during the launching and observing the resulting motion. The behavior of the model during these tests can be recorded using either the shadowgraph equipment already described or by using a high speed motion-picture camera.

A sample of the data which have been obtained in this facility is given in Fig. 10 which presents the drag of a 60° cone-cylinder model as a function of Mach number. The models were launched using a .220 Swift rifle and a Lucite sabot as shown in Fig. 6. The primary purpose of the sabot for these tests was to take the rifling of the gun and thus leave the model aerodynamically smooth. As shown in Fig. 10, data were obtained for Mach numbers from 1.5 to greater than 8. Over this range the experimental data deviate from the faired curve an average of about 0.8 percent and the maximum scatter is only three percent. Since the drag of this model is relatively high, it might be expected that the scatter would be greater for better aerodynamic shapes having less drag. However, we believe that with the foreseeable improvements in equipment and measuring techniques the scatter should not be significantly increased.

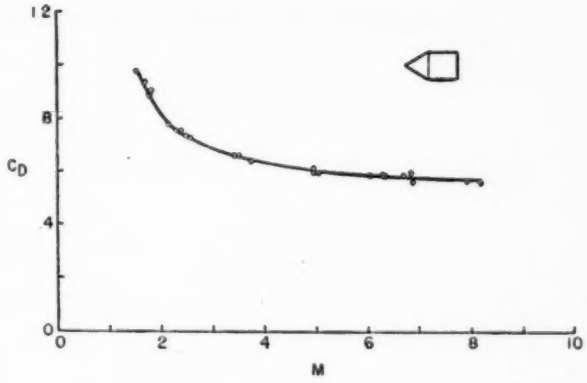


FIG. 10. Variation of drag with Mach number for 60° cone-cylinder model.

[3] A. C. Charters and R. A. Turetsky, "Determination of base pressure from free-flight data," Aberdeen Ballistic Research Laboratory (March, 1948).

The Application of Rate-Process Theory to Glass. I. Breaking Strength*

F. B. Hodgdon†
Department of Ceramics
D. A. Stuart‡ and F. E. Bjorklund
Department of Physics
University of Utah, Salt Lake City, Utah
(Received June 5, 1950)

An equation relating the ultimate strength of glass and porcelain to time under load is derived using the rate-process theory of Eyring. The data of Preston and co-workers are used to calculate the constants in the equation. The agreement between theoretical and experimental results is excellent.

I. INTRODUCTION

RECENT work of F. W. Preston and his co-workers[1–3] has provided data on the ultimate strength of glass and porcelain as a function of the time under load over a wide range of time. Because of the nature of these data it is possible to apply to ceramic materials the rate-process theory of plastic flow, developed originally by Eyring and others for textile fibers. Minor modifications of the mathematical treatment as presented by Eyring are necessary, probably because of the inherent difference between glass and fibers. These modifications do not affect the rigor of the mathematical derivation, but the physical reasons for them are still obscure and are the basis for further study.

Tobolsky and Eyring[4] have pointed out that in all unit processes of deformation, stress produces a two-fold effect on rate of strain. There is, first of all, an elastic component for which the rate of strain varies linearly with rate of stress and, secondly, a contribution to the rate of strain by relaxation processes. The latter can be readily measured for rubber, plastics, and metals, but data on ceramic materials are meager. Stress-strain, and other relationships can be derived in many ways, but it is simplest to derive them by the use of models in which the elastic component is represented by a pure spring and the relaxation process by a spring-dashpot combination.

II. PRESENT WORK

To analyze Preston's data, a model such as that illustrated in Fig. 1 was used. In this model the element shown as a spring-dashpot combination S_2-D represents a unit of the material which moves from one equilibrium position to another under the influence of an external applied force and in which bonds broken in one position are reformed in a successive position. The element shown as a pure spring, S_1 represents a network of bonds, probably primary, which once broken do not repair. (There is evidence that these bonds may repair at elevated temperatures; in this case, the pure spring should be replaced by a spring-dashpot combination.)

The spring S_1 has a modulus g_1, the spring S_2, a modulus g_2, and the dashpot constant is $\alpha=a\lambda_2/2kT$, where a is a cross-sectional area of the flow unit normal to the direction of the flow, λ_2 is the distance between equilibrium positions of the flow unit, k is the Boltzmann constant and T is the absolute temperature. A detailed discussion of this definition of α is given elsewhere.[5] In this equation $a\lambda_2$ is the volume of the hole available for flow.

When an external force f_0 is applied to the system, let the load on S_1 be f_1 and the load on S_2 and D be f_2. Let the elongation of S_1 be l_1, that of S_2, l_2 and that of D, l_3. After the load is applied and the initial elastic deformation of the system has taken place, the rate of flow of the dashpot is given by the equation

$$dl_3/dt=K\sinh\alpha f_2, \tag{1}$$

where K is the rate constant for the dashpot.[6] If the spring S_2, after the initial force has been applied, is stressed so that the force on it is f_{20}, then the force on S_2 at any time will be $f_{20}-g_2l_3$.

From this, by substitution in Eq. (1), we obtain

$$dl_3/dt=K\sinh\alpha(f_{20}-g_2l_3). \tag{2}$$

It is important to note at this point that Eq. (1) and the equation $f_2=f_{20}-g_2l_3$ are valid whether element 2 of the model is in compression or tension. Motion of the dashpot will always be in the direction of the applied force and this motion will decrease the force on the spring S_2.

If $f_{20}-g_2l_3$ is large, we can write, as a reasonably close

* This work was supported by the ONR, United States Navy Department, Research Contract No. N7 onr-451, Task Order 1, Project Number NR-032-168.
Presented at the Spring meeting of the American Ceramic Society, Glass Division, April, 1948.
† Now at American Lava Company, Chattanooga, Tennessee.
‡ Now at Department of Engineering Materials, Cornell University, Ithaca, New York.
[1] T. C. Baker and F. W. Preston, J. App. Phys. **17**, 162–170 (1946).
[2] T. C. Baker and F. W. Preston, J. App. Phys. **17**, 170–178 (1946).
[3] J. L. Glathart and F. W. Preston, J. App. Phys. **17**, 189–195 (1946).
[4] A. Tobolsky and H. Eyring, J. Chem. Phys. **11**, 125–134 (1943).
[5] The derivation of this equation can be found in Glasstone, Laidler, and Eyring, *Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, 1941). K is defined by
$$K=(2kT/h)\exp(-\Delta F^*/RT).$$

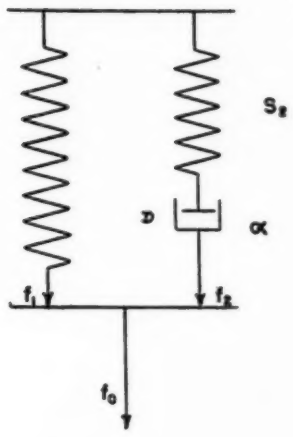


FIG. 1. Model used in derivations.

approximation

$$\frac{dl_3}{dt}=\frac{K\exp[\alpha(f_{20}-g_2l_3)]}{2}$$

which upon separation of the variables becomes

$$\frac{dl_3}{\exp[\alpha(f_{20}-g_2l_3)]}=\frac{Kdt}{2}. \tag{3}$$

Integration of Eq. (3) between the limits of $l_3=0$ when $t=0$ and $l_3=l_3$ when $t=t$ yields

$$(1/\alpha g_2)\exp[-\alpha(f_{20}-g_2l_3)]=Kt/2$$

which can be expressed in logarithmic form and arranged to give

$$f_{20}-g_2l_3=-(1/\alpha)\ln(\alpha g_2K/2)-(1/\alpha)\ln t. \tag{4}$$

Since $f_{20}-g_2l_3=f_2$ and $f_2=f_0\pm f_1$, Eq. (4) can be written

$$f_1=f_0+(1/\alpha)\ln(\alpha g_2K/2)+(1/\alpha)\ln t \text{ if } S \text{ is in tension} \tag{5a}$$

or

$$f_1=f_0-(1/\alpha)\ln(\alpha g_2K/2)-(1/\alpha)\ln t \text{ if } S \text{ is in compression.} \tag{5b}$$

This equation is equally valid with the proper sign when f_2 is a tension force and when it is a compression force. We shall consider f_2 to be a compression force and use Eq. (5b).

It has been shown by Tobolsky and Eyring[4] that with a statistical breaking of bonds, the rate of decrease of the number of primary bonds is given by the equation

$$-\frac{dN}{dt}=\frac{NkT}{h}\exp(-\Delta F^*/RT)2\sinh\frac{f_1\lambda_1}{2kTN} \tag{6}$$

where N is the number of primary bonds at time t, ΔF^* is the free energy of activation, k is the Boltzmann constant, h is Planck's constant, R is the gas constant, f_1 is the load on these bonds which causes an elongation λ_1, and T is the absolute temperature. Substitution of the value of f_1 given in Eq. (5a) in Eq. (6) gives

$$-dN/dt=(NkT/h)\times[\exp-(\Delta F_1^*/RT)]2\sinh\{[f_0-(1/\alpha)\times\ln(\alpha g_2K/2)-(1/\alpha)\ln t]\lambda_1/2kTN\}. \tag{7}$$

If the bonds do not repair after breaking, Eq. (7) may be written as an exponential:

$$-dN/dt=(NkT/h)\exp\left\{-\frac{\Delta F_1^*}{RT}+\frac{[f_0-(1/\alpha)\ln(\alpha g_2K/2)-(1/\alpha)\ln t]\lambda_1}{2kTN}\right\}. \tag{8}$$

Substituting $c=(1/\alpha)\ln(\alpha g_2K/2)$ in Eq. 8 gives

$$-dN/dt=(NkT/h)\exp(-\Delta F_1^*/RT)\times\exp[(f_0-c)\lambda_1/2kTN]t^{(-\lambda_1/2kNTa)}. \tag{9}$$

Equation (9) cannot be integrated unless further assumptions are made. In the definition $\alpha=a\lambda_2/2kT$, the numerator, $a\lambda_2$, is the average volume of the unit, the bonds of which break and repair. It is presented in the model by the dashpot. For the other bonds, those which do not repair after breaking and which are represented in the model by the pure springs, the volume is λ_1/N. As an approximation, we assume that the ratio of the volume of the non-repairable bonds to the volume of the repairable bonds is substantually constant until fracture occurs. With this assumption Eq. (9) becomes:

$$-dN/dt=(NkT/h)\exp-(\Delta F_1^*/RT)\times\exp[(f_0-e)\lambda_1/2kTN]t^{-(v_1/v_2)} \tag{10a}$$

which is equivalent to

$$-\frac{dN}{(NkT/h)\exp-(\Delta F_1^*/RT)\times\exp[(f_0-c)\lambda_1/2kTN]}=t^{-(v_1/v_2)}dt. \tag{10b}$$

When $t=0$, $N=N_0$ and when $t=t_1$, $N=0$, $v_1=\lambda_1/N_0$ and $v_2=a\lambda_2$. Equation (10b) has been integrated by Tobolsky and Eyring and gives

$$\gamma\ln t=-(f_0-c)\lambda/2kTN_0-\ln(f_0-c)\lambda/2kTN_0+(\Delta F^*-RT\ln\gamma)/RT. \tag{11}$$

III. EFFECT OF TEMPERATURE UPON BREAKING STRENGTH

If a constant load be applied to a glass specimen, it is well known that the apparent strength depends upon the temperatures. That is to say that if a constant load is applied to a specimen of glass at various temperatures the glass requires different times to fracture. Further if the time required to fracture a specimen of glass at constant load be measured and plotted *versus* the temperature, then a minimum is observed in the plotted curve.

We shall assume that v_1 and v_2 are approximately

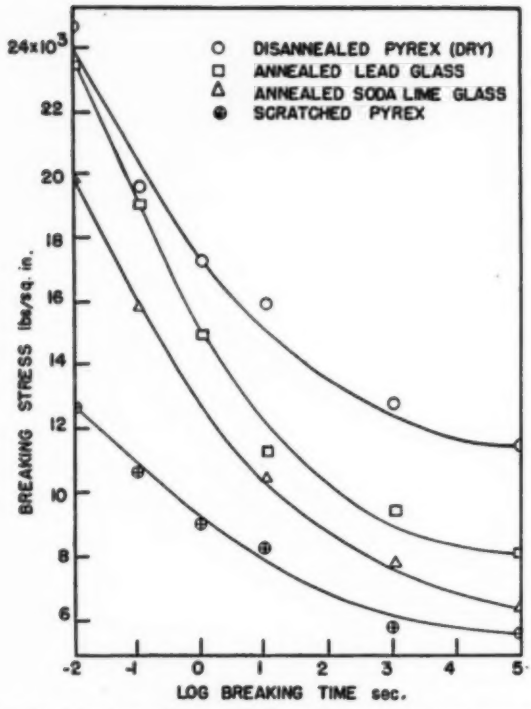


FIG. 2. Comparison of theory and experiment, breaking strength *vs.* temperature for certain glasses (after Glathart and Preston).

independent of temperature, and that $\Delta F^* = \Delta H^* - T\Delta S^*$ where ΔH^* and ΔS^* are not temperature dependant. If a minimum occurs in the time-temperature relationship, the temperature at which this minimum occurs is that temperature at which $dt/dT = 0$. Hence, differentiating Eq. (11) with respect to temperature and setting $dt/dT = 0$, we find

$$0 = -\frac{\gamma}{t}\frac{dt}{dT} = \frac{(f_0 - c)v_1}{2kT}\left(\frac{1}{T} + \frac{1}{f_0 - c}\right)\frac{dc}{dt} + \frac{1}{f_0 - c}\frac{dc}{dT} - \frac{\Delta H_1^*}{RT^2} \quad (12a)$$

or

$$\frac{(f_0 - c)v_1}{2kT^2} + \left(\frac{v_1}{2kT} + \frac{1}{f_0 - c}\right)\frac{dc}{dT} - \frac{\Delta H_1^*}{RT^2} = 0 \quad (12b)$$

now

$$c = (1/\alpha)\ln(\alpha g_2 K/2)$$

or

$$c = (2kT/v_2)\ln(v_2 g_2/2h) - (2k\Delta H_2^*/Rv_2) + (2kT\Delta S_2^*/Rv_2) \quad (13)$$

and

$$dc/dT = (2k/v_2)\ln(v_2 g_2/2h) + (2k\Delta S_2^*/Rv_2). \quad (14)$$

Thus after substituting the value of dc/dT from Eq. (14) into Eq. (12a) and multiplying through by $(f_0 - c)2kT^2$ we find

$$(f_0 - c)(v_1 f_0 + (2k\Delta H_1^*\gamma/R)) + 2kT[c + (2k\Delta H_2^*/Rv_2)] = 0 \quad (15)$$

or substituting for c and rearranging

$$T^2 - \left(\frac{v_1 f_0}{2k} + \frac{2\Delta H_1^*\gamma}{R}\right)T + [(v_1 f_0/2k + \Delta H_1^*\gamma/R) \times (f_0 v_1/2k + \Delta H_2^*/R)/(\ln(v_2 g_2/2h) + \Delta S_2^*/R)] = 0.$$

Solving for T by the quadratic formula, we find,

$$T = \frac{v_1 f_0}{4k} + \frac{\Delta H_1^*\gamma}{R} \pm \left\{\left(\frac{v_1 f_0}{4k} + \frac{\Delta H_1^*\gamma}{R}\right)^2 - [(v_1 f_0/2k + \Delta H_1^*/R)(f_0 v_2/2k + \Delta H_2^*/R)/(\ln(v_2 g_2/2h) + \Delta S_2^*/R)]\right\}^{\frac{1}{2}}. \quad (16)$$

It is obvious that at both the above temperatures we cannot possibly observe minima. To find at which temperature a minimum occurs, we solve for d^2t/dT^2 and substitute the above temperatures. If $d^2t/dT^2 > 0$, then at that point we observe a minimum. When this is done we find that at the lower temperature we have a minimum and at the higher a maximum.

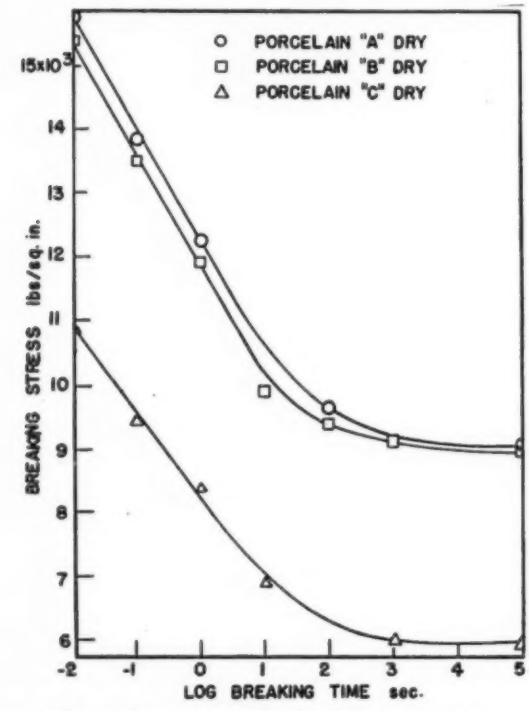


FIG. 3. Comparison of theory and experiment, breaking strength *vs.* temperature for certain porcelains (after Glathart and Preston).

IV. RESULTS AND DISCUSSION

The data of Preston have been used to calculate the parameters in Eq. (10a) and Figs. 2 and 3 show the results. The smooth curves are plots of the equation

TABLE I. Values of constants in the breaking strength equation.

Material	γ	ΔF_1^* kilocalories	λ_1/N_0 in^3	c psi
Disannealed Pyrex (dry)	0.272	17.1*	5.56×10^{-24}	11,000
Scratched Pyrex (wet)	0.271	17.2*	11.4×10^{-24}	5,500
Annealed lead (wet)	0.375	17.6*	7.3×10^{-24}	7,500
Annealed soda lime (wet)	0.239	16.9*	4.08×10^{-24}	5,700
Porcelain A	1.000	19.9	78.9×10^{-24}	9,100
Porcelain B	1.000	19.9	78.3×10^{-24}	8,850
Porcelain C	1.000	19.6	98.9×10^{-24}	8,150

* ΔF_1^* as tabulated is the apparent ΔF_1^* and should be increased by $0.5961n(1/\gamma)$.

when the values listed in Table I are used, and it appears that the breaking equation fits the data rather well. Further, the values obtained for the free energy of activation appear reasonable, since they are of the same order of magnitude as the strengths of alkali-oxygen and alkaline earth bond strengths in glass as determined by other methods.[6] However, as previously pointed out, a model was used in which a dashpot associated with spring S_1 has been neglected because the constant of this dashpot is such that it does affect the results at temperatures below 300°C, approximately. The simplified model presented here cannot be used in analyses of the temperature dependence of properties such as viscosity and birefringence which are now under way in these laboratories.

Data were not available with which to check reliably the breaking strength temperature results. However, calculations on a typical glass yield approximately 300°K for the minimum and 4000°K for the maximum strengths, respectively. Obviously the second temperature is so high that the maximum would never be observed.

It was assumed in our derivation that there was only an initial elongation when a load was applied, which is in accord with the known absence of "creep" in glass at ordinary temperatures; the theoretical admissability of this assumption will depend upon the slope of potential energy barrier for glass forming oxides.

The breaking equation predicts that failures in glass will be failures in tension, which is, we believe, in accord with Preston's views.

The brittle nature of glass can be explained in a qualitative manner by these results. If the strength of the alkali-oxygen and alkaline earth-oxygen bonds is of the same order of magnitude as the free energy of activation of the viscous element bond, a force sufficient to start viscous flow will also be sufficient to rupture these primary bonds. In other words, a force sufficient to cause flow, which we believe arises from the presence of "holes" in the structure is large enough to rupture bonds which do not repair. The nature of these "holes" is at present the subject of further investigation. This condition will be altered somewhat at elevated temperatures when the influence of the dashpot on spring S_1, neglected here, becomes appreciable.

ACKNOWLEDGMENT

The authors wish to express their appreciation to the staff of the School of Mines and Mineral Industries, and to H. Eyring, Dean of the Graduate School, University of Utah, for their many suggestions which were invaluable in the preparation of this paper.

[6] Kuan-Ham Sun, J. Am. Cer. Soc. **30**, 277–281 (1947).

The Application of Rate-Process Theory to Glass. II. Viscosity*

F. B. Hodgdon†
Department of Ceramics
D. A. Stuart‡
Department of Physics
University of Utah, Salt Lake City, Utah
(Received June 5, 1950)

In previous efforts to derive an equation for the temperature-viscosity relation of glass it has been assumed, tacitly or otherwise, that glass is a simple liquid. Since glass is not a simple liquid, it is necessary to find a model that will, in its reaction to applied stresses, have the same characteristics as a specimen of glass. By the use of Eyring's rate-process theory, such a model is proposed and several equations relating viscosity to time and to temperature are derived. The equations include an equation that relates the viscosity of any glass specimen to time, one which defines the rate of removal of internal strain, and an equation which relates the viscosity of a specimen an infinite time after beginning of a test to the temperature of the test. This final viscosity is the viscosity usually reported in the literature. It is found that the parameters in the equation reduced to infinite time can be calculated from the chemical composition of the glass and the viscosity of any given glass can be predicted at any temperature over a very wide range. Curves and tables are included to show the agreement of experiment and theoretical values.

I. INTRODUCTION

THE equation most generally used to define the viscosity-temperature relation of glass was derived by Sheppard[1] on the basis of the Maxwell-Boltzmann Distribution law. Over a limited temperature range, usually not exceeding 300°C, the deviations between observed and calculated viscosities are not large. Over extended temperature ranges, however, such as from 900°K to 1900°K, the agreement between theoretical and experimental values is not good. In fact, with some glasses the deviation is as much as tenfold, the reason for which is readily apparent. Implicit in Sheppard's derivation is the assumption that glasses are simple liquids. In view of the extensive literature indicating that glasses are not simple liquids, it is not surprising that an equation based upon the assumption that they are does not yield results in agreement with experiment.

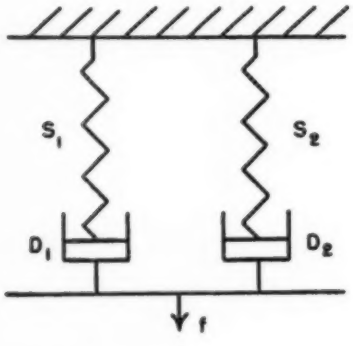


FIG. 1. Model used in derivations.

In the temperature range where the viscosity of glass is greater than 10^{10} poises, viscosities are usually measured experimentally by measuring the rate of elongation of a glass rod that is held at constant temperature and is subjected to a constant load. The quotient of the stress and the rate of elongation is, by definition, the viscosity. Data so taken can be handled readily by the type of mathematical analysis that was employed in a previous paper[2] from these laboratories to derive an equation relating time under load to the ultimate strength of glass. In that paper a model was used which consisted of pure springs and dashpots so joined that the behavior of the model under the influence of an external load represented the behavior of a glass specimen under load. It was found that a model consisting of two spring-dashpot combinations in parallel permitted the derivation of a theoretical equation for breaking strength which gave values in close agreement with experiment. The mathematical treatment of the problem was simplified by neglecting one of the dashpots; this was permissible because the characteristic of the neglected dashpot was such that it did not affect results at room temperature.

It was pointed out in that paper, however, that since the characteristic of this neglected dashpot is temperature dependent, the simplification would probably be inadmissible at higher temperatures, such as those at which the viscosity of glass is usually measured. This has been found to be the case, and in the work on viscosity, herein reported, both springs and both dashpots have been used in the mathematical derivations.

The model shown in Fig. 1 represents a glass rod subjected to a constant, external, tensile stress; since the viscosity of a given glass, as measured by a method that applies a constant tensile stress to a glass rod,

* This work was supported by the ONR, United States Navy Department, Research Contract No. N7 onr-451, Task Order 1. Project Number NR-032-168.
Presented at the Spring Meeting of the American Ceramic Society, Glass Division, April, 1949.
† Now at American Lava Company, Chattanooga, Tennessee.
‡ Now at Department of Engineering Materials, Cornell University, Ithaca, New York.
[1] S. E. Sheppard, Nature **125**, 489, 709 (1930).
[2] Hodgdon, Stuart, and Bjorklund, J. App. Phys. **21**, 1156 (1950).

maintained at a constant temperature, does not differ from the viscosity at that temperature, as measured by other methods, the mathematical analysis applied to this model should yield results which are applicable, regardless of the experimental method employed. It has been shown by Lillie[3] and by Dale and Stanworth[4] that the viscosity of a given glass depends, among other things, on its thermal history; a satisfactory approach to the problem of viscosity should therefore consider, first of all, the manner in which a glass specimen, maintained at constant temperature, will approach an equilibrium value for viscosity with time. This is the value for viscosity, of course, most generally reported in the literature. Therefore, the viscosity-time relation at constant temperature for a specimen initially free from internal strain, will be considered first. This is, mathematically, the simplest of the time-temperature-viscosity relations. Such freedom from strain can be attained in the laboratory by maintaining the specimen for a long period of time at a temperature very near to that at which viscosity measurements are to be made, as was done by Lillie.

The second relation to be considered will be that of the rate of removal of strain initially present in a specimen. In this case, it will not be possible to verify the rate equation which will be derived by direct comparison of predicted and experimental values, as is usually done, because there is no known way to measure, quantitatively, the amount of strain present in a sample at any time. It is true that the amount of strain may be estimated by the use of polarized light but the assumption that the amount of strain is proportional to the birefringence[5] has never been established.

The third and final relation to be analyzed is the temperature dependence of viscosity under all conditions of initial internal strain. This is the most difficult, mathematically, of the time-temperature-viscosity relations and requires the equation for the rate of strain removal for its solution. It will be possible to check the validity of the assumptions which will have to be made to derive this equation in two ways. First of all, by assuming that the initial strain is zero, it should be possible to reduce the general equation to the one derived for the special case of no initial strain and as a second check, a comparison between observed and predicted values of viscosity over a wide temperature range can be made. It will be shown that equilibrium viscosities are independent of the amount or nature of initial strain present in the specimen.

It will finally be shown that it is possible to correlate the values of the parameters in the equilibrium viscosity-time equation with the chemical composition of the glass. It will be shown that it is possible, given the chemical composition of the glass as determined by analysis, to predict the equilibrium viscosity at any temperature.

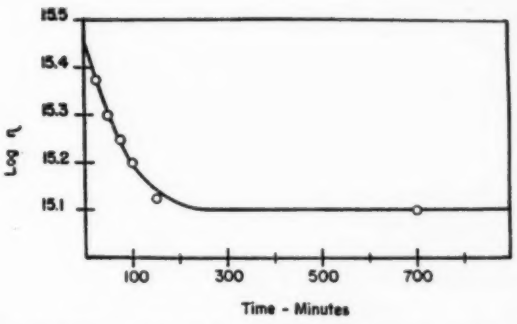


FIG. 2. Comparison of theory and experiment, log viscosity *vs.* time, for annealed glass at 486.7°C (after Lillie, see reference 3).

Two limitations are placed upon the comparison of calculated and observed values for viscosity at any temperature: the specimen must be at uniform temperature throughout its length and the viscosity determinations must be made under conditions of constant temperature. That these conditions exist is implicit in the derivations of the time-temperature-viscosity relations and represents, in any case, only those experimental conditions which would yield reproducible results.

II. NOTATION

The model used in the mathematical analyses is shown in Fig. 1. The springs represent those bonds in the glass which are stretched to their elastic limits under the influence of an external applied force and then break. They do not reform in any manner. The dashpots represent those bonds which break and then reform under the influence of the same applied force. It is not necessary that they reform in the same manner or position as that in which they originally existed; in fact, it is highly improbable that they do. No assumptions are made regarding the nature of the forces which constitute these bonds.

It is a fundamental assumption of Eyring's rate-process theory that all flow processes, of which viscous flow is one, proceed by a series of "jumps," in which discreet units of molecular dimensions move from one equilibrium position to another under the influence of an external applied force. This motion is analogous, mathematically, to the motion of a dashpot. No attempt is made to differentiate between dashpots; they differ only in that they have different temperature dependent characteristics; that is, there will be a difference in the change in their rates of flow with temperature. Both dashpots, however, obey the hyperbolic sine law; that is, their rates of flow are proportional to the hyperbolic sine of the applied force. A complete discussion of these dashpots is given elsewhere.[6]

[3] H. R. Lillie, J. Am. Cer. Soc. **16**, 619–631 (1933).
[4] A. E. Dale and J. E. Stanworth, J. Soc. Glass Tech. **29**, 414–427 (1945).
[5] G. W. Morey, *Properties of Glass* (Reinhold Publishing Corporation, New York, 1938), pp. 342ff.
[6] Glasstone, Ladler, and Eyring, *Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, 1941).

The following symbols are used throughout this paper.

f = the total external applied force per unit area.
f_1 = the total stress on unit 1.
f_2 = the total stress on unit 2.
$f_{1,0}$ = the total stress on unit 1 at $t=0$.
$f_{2,0}$ = the total stress on unit 2 at $t=0$.
f_1' = the stress on unit 1 due to f.
f_2' = the stress on unit 2 due to f.
f'' = the total internal stress.
f_1'' = the stress on unit 1 due to f''.
f_2'' = the stress on unit 2 due to f''.
l = the total strain (elongation per unit length).
l_1 = the total elongation of spring S_1 at any time.
l_2 = the total elongation of spring S_2 at any time.
l_1' = the elongation of spring S_1 at any time caused by f_1'.
l_2' = the elongation of spring S_2 at any time caused by f_2'.
l_1'' = the elongation of spring S_1 at any time caused by f''.
l_2'' = the elongation of spring S_2 at any time caused by f''.
g_1 = the spring constant of spring S_1 so that $f_1=g_1l_1$.
g_2 = the spring constant of spring S_2 so that $f_2=g_2l_2$.
k = Boltzmann constant—1.37×10^{-16} ergs per degree.
h = Planck's constant—6.624×10^{-27} erg seconds.
R = gas constant—1.988 calories per degree.
T = temperature in degrees Kelvin.
η = viscosity in poises.
t = time in seconds.
ΔF_1^* = free energy of activation of unit 1 in calories per mole.
ΔF_2^* = free energy of activation of unit 2 in calories per mole.
ΔH_1^* = heat of activation of unit 1 in calories per mole at absolute zero.
ΔH_2^* = heat of activation of unit 2 in calories per mole at absolute zero.
ΔS_1^* = entropy of activation of unit 1 in calories per degree.
ΔS_2^* = entropy of activation of unit 2 in calories per degree.
α_1 = "constant" of dashpot 1.
α_2 = "constant" of dashpot 2.
K_1 = "characteristic" of dashpot 1.
K_2 = "characteristic" of dashpot 2.
v_1 = volume of the "hole" necessary for flow of unit 1 in cubic centimeters.
v_2 = volume of the "hole" necessary for flow of unit 2 in cubic centimeters.
λ_1 = distance in the direction of flow between equilibrium positions of unit 1 in centimeters.
λ_2 = distance in the direction of flow between equilibrium positions of unit 2 in centimeters.
a_1 = cross-sectional area of the "hole" necessary for flow of unit 1 in square centimeters.
a_2 = cross-sectional area of the "hole" necessary for flow of unit 2 in square centimeters.

The following mathematical relations exist:

$$\alpha=v/2kT,$$
$$v=a\lambda,$$

and

$$K=(2kT/h)\exp(-\Delta F^*/RT).$$

III. VISCOSITY-TIME RELATION IN THE ABSENCE OF INTERNAL STRAIN

In Fig. 1, S_1 and S_2 are pure springs, to each of which is connected, in series, a dashpot of the kind proposed by Eyring and others; that is, the dashpots are such that their rates of flow are proportional to the hyperbolic sine of the applied force:

$$dl/dt=K\sinh\alpha f.$$

At any time t the total rate of elongation will be the sum of the rate of elongation of each spring and that of its associated dashpot,

$$dl/dt=dl_1/dt+K_1\sinh\alpha_1f_1, \tag{1a}$$

$$dl/dt=dl_2/dt+K_2\sinh\alpha_2f_2. \tag{1b}$$

Since, by Hooke's law, $f_1=g_1l_1$ and $f_2=g_2l_2$, it follows that:

$$dl_1/dt=(1/g_1)(df_1/dt) \tag{2a}$$

and

$$dl_2/dt=(1/g_2)(df_2/dt). \tag{2b}$$

Since $f_1+f_2=f$, the total external applied force per unit area, which is constant:

$$df_1/dt=-df_2/dt. \tag{3}$$

Thus after elimination of f_1, l_2, and l_1, Eqs. (1a) and (1b) become:

$$dl/dt=-(1/g_1)(df_2/dt)+K_1\sinh\alpha_1(f-f_2)$$
$$=(1/g_2)(df_2/dt)+K_2\sinh\alpha_2f_2. \tag{4}$$

If $\alpha_1(f-f_2)$ and α_2f_2 are small, as a reasonably close approximation, Eq. (4) may be written:

$$dl/dt=-(1/g_1)(df_2/dt)+K_1\alpha_1(f-f_2)$$
$$=(1/g_2)(df_2/dt)+K_2\alpha_2f_2. \tag{5}$$

If

$$m=K_1\alpha_1f,$$
$$n=K_1\alpha_1+K_2\alpha_2,$$
$$p=g_1g_2/(g_1+g_2),$$

Eq. (5) becomes

$$df_2/dt=p(m-nf_2) \tag{6}$$

which has the solution

$$f_2=m/n-De^{-npt}. \tag{7}$$

In a previous work[2] it was assumed that when the external load was applied, an immediate elastic deformation occurred and that further elongation of the model was caused by movement of the dashpots. The force on the spring, before the dashpots had started to move, was called $f_{2,0}$. Using this notation we have, at $t=0$ and $f_2=f_{2,0}$

$$D=m/n-f_{2,0}=[K_1\alpha_1f/(K_1\alpha_1+K_2\alpha_2)]-f_{2,0} \tag{8}$$

by means of which the constant of integration can be evaluated. Since $dl/dt=(1/g_2)(df_2/dt)+K_2\alpha_2f_2$ [Eq. (4)], we obtain, by substitution of the value of f_2 in Eq. (7)

$$dl/dt=K_2\alpha_2m/n+D(np/g_2-K_2\alpha_2)e^{-npt}. \tag{9}$$

Substitution of the previously assigned values of m, n, and p, in Eq. (9) gives

$$dl/dt=K_1\alpha_1K_2\alpha_2f/(K_1\alpha_1+K_2\alpha_2)$$
$$+D\{[(K_1\alpha_1+K_2\alpha_2)g_1-(g_1+g_2)K_2\alpha_2]/(g_1+g_2)\}$$
$$\times\exp[-(K_1\alpha_1+K_2\alpha_2)g_1g_2t/(g_1+g_2)].$$

Since, by definition the viscosity η is $f/dl/dt$, it follows that

$$\eta=\frac{1}{K_1\alpha_1K_2\alpha_2/(K_1\alpha_1+K_2\alpha_2)+D/f[(K_1\alpha_1g_1-K_2\alpha_2g_2)/(g_1+g_2)]\exp[-(K_1\alpha_1+K_2\alpha_2)g_1g_2t/(g_1+g_2)]}. \quad (11)$$

At equilibrium, when $t=\infty$, Eq. (11) reduces to

$$\eta=(K_1\alpha_1+K_2\alpha_2)/K_1\alpha_1K_2\alpha_2. \quad (12)$$

By the use of the equations defining K, α, and v, Eq. (12) becomes

$$\eta=[(\exp(\Delta F_1^*/RT))/v_1+(\exp(\Delta F_2^*/RT))/v_2]. \quad (13)$$

In his work on viscosity, Lillie[3] held a glass at 477.8°C for 64 hours before making measurements. From his data, the following values for the parameters in Eq. (17) have been calculated

$$K_1\alpha_1K_2\alpha_2/(K_1\alpha_1+K_2\alpha_2)=8\times10^{-16} \text{ reciprocal poises,}$$
$$(D/f)[(K_1\alpha_1g_1-K_2\alpha_2g_2)/(g_1+g_2)]=5.2\times10^{-16} \text{ reciprocal poises,}$$
$$[(K_2\alpha_1+K_2\alpha_2)g_1g_2/(g_1+g_2)]=1.66\times10^{-2} \text{ reciprocal seconds.}$$

With these empirical values of the parameters, the viscosity-time curve which is shown in Fig. 2 was plotted. The smooth curve represents calculated values; the points represent Lillie's experimental observations.

IV. RATE OF STRAIN REMOVAL

Assume that the springs of the model shown in Fig. 1 are stressed in some way before they are connected to their respective dashpots. No assumptions are made regarding the nature of these stresses or the manner of their origin. Further, no assumption is made that the stress in each spring is the same. These stresses will, according to Hooke's law, satisfy the equations

$$f_1''=g_1l_1'' \quad (14a)$$

and

$$f_2''=g_2l_2''. \quad (14b)$$

The total rate of elongation of the system will be given by the equation:

$$dl''/dt=dl_1''/dt+K_1\sinh\alpha_1f_1''=dl_2''/dt+K_2\sinh\alpha_2f_2''. \quad (15)$$

Since $f''=f_1''+f_2''=g_1l_1''+g_2l_2''$, Eq. (15) can be written

$$df''/dt=g_1(dl_1''/dt)+g_2(dl_2''/dt)=g_1(dl''/dt)-g_1K_1\sinh\alpha_1f_1''+g_2(dl''/dt)-g_2K_2\sinh\alpha_2f_2''. \quad (16)$$

If the internal stresses are small, that is, if the glass is not badly strained, dl''/dt will be negligible; it will be true also, as a reasonably close approximation, that the hyperbolic sine is equal to its argument. With these approximations Eq. (16) becomes:

$$df''/dt=-g_1K_1\alpha_1f_1''-g_2K_2\alpha_2f_2''. \quad (17)$$

A solution of this differential equation is

$$f_1''=(\beta_1/\alpha_1)e^{-g_1K_1\alpha_1t}, \quad (18a)$$

$$f_2''=(\beta_2/\alpha_2)e^{-g_2K_2\alpha_2t} \quad (18b)$$

where β_1 and β_2 are dimensionless constants of integration. Since $f''=f_1''+f_2''$, it follows that

$$f''=(\beta_1/\alpha_1)e^{-g_1K_1\alpha_1t}+(\beta_2/\alpha_2)e^{-g_2K_2\alpha_2t}. \quad (19)$$

If the internal forces are not small, a solution of Eq. (17) is

$$f''=(g_1+g_2)l''+(2/\alpha_1)\tanh^{-1}R_1e^{-\alpha_1g_1K_1t}+(2/\alpha_2)\tanh^{-1}R_2e^{-\alpha_2g_2K_2t}+C \quad (20)$$

where R_1, R_2 and C are constants. Equation (20) will reduce to Eq. (19), when l'' is approximately zero and $e^{-\alpha_1g_1K_1t}$ and $e^{-\alpha_2g_2K_2t}$ are small.

The above solutions to Eq. (17) are not general. They are, however, valid and sufficient for the present purpose; a complete analysis of this problem and its applications to annealing will be presented in a subsequent paper.

V. VISCOSITY-TIME RELATION IN THE PRESENCE OF INTERNAL STRAIN

Assume that strain is present in the specimen when viscosity measurements are started. This condition will be realized experimentally if measurements are made on a freshly drawn fibre, as was done by Lillie[3] and by Dale and Stanworth.[4]

When internal strain is present, the total stress on each unit of the model shown in Fig. 1 is the sum of the internal stress on that unit and that part of the external load taken by the unit. That is

$$f_1=f_1'+f_1'' \quad (21a)$$

and

$$f_2=f_2'+f_2''. \quad (21b)$$

Since $f=f_1'+f_2'$, and at constant load, $df/dt=0$, it follows that

$$df_1'/dt=-df_2'/dt. \quad (22)$$

From Eqs. (1a) and (1b):

$$dl/dt=dl_1/dt+K_1\sinh\alpha_1f_1=dl_2/dt+K_2\sinh\alpha_2f_2,$$

or, as a reasonably close approximation, if the forces are small:

$$dl/dt=(1/g_1)(df_1/dt)+K_1\alpha_1f_1=(1/g_2)(df_2/dt)+K_2\alpha_2f_2.$$

Since $df_2/dt=(df''/dt)-(df_1/dt)$, it follows that

$$(df_1/dt)(1/g_1+1/g_2)+f_1(K_1\alpha_1+K_2\alpha_2)=(l/g_2)(df''/dt)+K_2\alpha_2f''+K_2\alpha_2f. \quad (23)$$

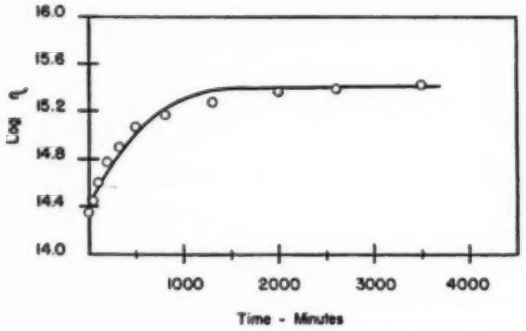


FIG. 3. Comparison of theory and experiment, log viscosity *vs.* time, for freshly drawn glass at 477.8°C (after Lillie, see reference 3).

It was shown in Eq. (19) that

$$f''=(\beta_1/\alpha_1)e^{-\alpha_1 g_1 K_1 t}+(\beta_2/\alpha_2)e^{-\alpha_2 g_2 K_2 t}.$$

Hence

$$df''/dt=-\alpha_1 g_1 K_1 f_1''-\alpha_2 g_2 K_2 f_2''$$

from which it follows that

$$df_1/dt+f_1 g_1 g_2(K_1\alpha_1+K_2\alpha_2)/(g_1+g_2)$$
$$=K_2\alpha_2 g_1 g_2 f/(g_1+g_2)+[g_1 g_2/(g_1+g_2)]$$
$$\times(K_2\alpha_2-(g_1/g_2)K_1\alpha_1)(\beta_1/\alpha_1)e^{-\alpha_1 g_1 K_1 t}. \quad (24)$$

If

$$m=(K_1\alpha_1+K_2\alpha_2)g_1 g_2/(g_1+g_2),$$
$$n=K_2\alpha_2 g_1 g_2 f/(g_1+g_2),$$
$$p=[g_1 g_2/(g_1+g_2)][K_2\alpha_2-(g_1/g_2)K_1\alpha_1](\beta_1/\alpha_1),$$

Eq. (24) can be written

$$df_1/dt+mf_1=n+pe^{-\alpha_1 g_1 K_1 t}. \quad (25)$$

For which a solution is

$$f_1=n/m+[p/(m-\alpha_1 K_1 g_1)]e^{-\alpha_1 g_1 K_1 t}+Ce^{-mt}. \quad (26)$$

When $t=0$, $f_1=f_{1,0}$ and $f_2=f_{2,0}$, where $f_{1,0}$ and $f_{2,0}$ have the meanings previously given. By means of these boundary conditions it is possible to evaluate the constant of integration in Eq. (26):

$$C=f_{1,0}-\frac{K_2\alpha_2 f}{(K_1\alpha_1+K_2\alpha_2)}-[(g_2K_2\alpha_2+g_1K_1\alpha_1)\beta_1/$$
$$(K_2\alpha_2 g_2 f-K_1\alpha_1 g_1-K_1\alpha_1 g_2)\alpha_1]. \quad (27)$$

Differentiation of Eq. (26) with respect to time gives

$$df_1/dt=-[p\alpha_1 K_1 g_1/(m-\alpha_1 K_1 g_1)]$$
$$\times e^{-\alpha_1 g_1 K_1 t}-mCe^{-mt}. \quad (28)$$

Since $dl/dt=+(1/g_1)(df_1/dt)+K_1\alpha_1 f_1$, [Eq. (5)], Eq. (28) can be written in the form:

$$dl/dt=pe^{-\alpha_1 g_1 K_1 t}$$
$$\cdot[(K_1\alpha_1-\alpha_1 g_1 K_1/g_1)/(m-\alpha_1 K_1 g_1)]$$
$$+e^{-mt}(-mC/g_1+K_1\alpha_1 C)+K_1\alpha_1 n/m. \quad (29)$$

Substitution of the values of m, n, and p, previously assigned, in Eq. (29) and simplifying gives

$$dl/dt=\frac{K_1\alpha_1 K_2\alpha_2 f}{(K_1\alpha_1+K_2\alpha_2)}+\left[\frac{C(K_1\alpha_1 g_1+K_2\alpha_2 g_2)}{(g_1+g_2)}\right]$$
$$\times\exp\left[\frac{-(K_1\alpha_1+K_2\alpha_2)g_1 g_2 t}{(g_1+g_2)}\right]. \quad (30)$$

A comparison of Eq. (30) with Eq. (10) shows that it differs only in the constant C: If the initial strain is zero, $\beta_1=0$ and

$$C=f_{1,0}-K_2\alpha_2 f/(K_1\alpha_1+K_2\alpha_2).$$

Since $f_{1,0}=f-f_{2,0}$, we can write

$$C=K_1\alpha_1 f/(K_1\alpha_1+K_2\alpha_2)-f_{2,0}$$

which is identical with the value of D in Eq. (8). At infinite time, Eq. (30) will reduce to Eq. (13).

$$\eta=h[\exp(\Delta F_1^*/RT)/v_1+\exp(\Delta F_1^*/RT)/v_2].$$

The values of the parameters in Eq. (30) for glasses having internal strain have been calculated from the data of Lillie[3] and of Dale and Stanworth.[4] With these values which are listed in Table IV, the viscosity-time relation of the glasses used have been calculated. The results are shown in Figs. 3 to 5 inclusive. The smooth curves represent theoretical values; the points are observed values.

VI. RELATION OF VISCOSITY TO COMPOSITION

An examination of the general viscosity equation [Eq. (19)] shows that it contains four temperature dependent parameters ΔF_1^*, ΔF_2^*, v_1, and v_2. In order to correlate viscosity and composition, it was necessary to establish the form of the function for each of these parameters. This cannot be done in an exact mathematical manner in the present state of the knowledge of glass structure, and it was therefore necessary to rely upon purely empirical relations. The relations which were finally adopted probably do not have any real significance; the best justification for them is that their use permits the calculation of the viscosity of a glass to a high degree of accuracy over an extended temperature range.

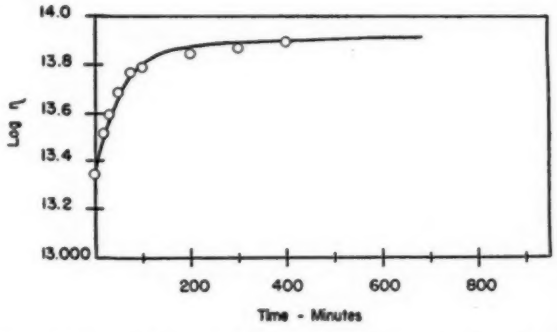


FIG. 4. Comparison of theory and experiment, log viscosity *vs.* time, for freshly drawn glass at 505.5°C (after Lillie, see reference 3).

It was first assumed that the free energies of activation were defined by the equations:

$$\Delta F_1^* = \Delta H_1^* - T\Delta S_1^*,$$
$$\Delta F_2^* = \Delta H_2^* - T\Delta S_2^*$$

and that ΔH_1^* and ΔH_2^* were independent of the temperature. In other words, it was assumed that the free energies of activation were linear functions of temperature and that the changes in them with temperature could be considered entropy changes.

A satisfactory form of the function for v_1 and v_2 was more difficult, but repeated trials finally led to the definitions:

$$v_1 = b_1 \exp(-d_1T + m_1T^2),$$
$$v_2 = b_2 \exp(-d_2T + m_2T^2)$$

where b_1, b_2 d_1, d_2, m_1, and m_2 are constants. However, it was found that unless the temperature T was very high the T^2 terms can be neglected. It was found that m_1 was of the order of magnitude of -2×10^{-8}.

Substitution of the above empirical functions in the viscosity equation, neglecting the terms in T^2, gives

$$\eta = h[\exp(\Delta H_1^*/RT - \Delta S_1^*/R + d_1T - \ln b_1) + \exp(\Delta H_2^*/RT - \Delta S_2^*/R + d_2T - \ln b_2)]$$

from which it follows, after factoring out $(\exp\Delta F_1^*/RT)/v_1$, that

$$\log\eta = (\Delta H_1^* - \Delta S_1^*T)/2.3RT + d_1T/2.3 + \log h/b_1 + \log\{1+\exp[(\Delta H_2^* - \Delta H_1^* + (\Delta S_1^* - \Delta S_2^*)T)/RT + (d_2-d_1)T + \ln b_1/b_2]\}. \quad (31)$$

If

$$M = \exp\{[\Delta H_2^* - \Delta H_1^* + (\Delta S_1^* - \Delta S_2^*)T]/RT + (d_2-d_1)T + \ln b_1/b_2\}$$

the viscosity equation can be written:

$$\log\eta = \log h + \Delta H_1^*/2.3RT - (\Delta S_1^*/2.3R + \log b_1) + d_1T/2.3 + \log(M+1) \quad (32)$$

where M has the meaning given above.

At temperatures where ΔF_1^* is much greater than ΔF_2^*, M becomes negligible and the viscosity equation reduces to

$$\log\eta = \log h + \Delta H_1^*/2.3RT - (\Delta S_1^*/2.3R + \log b_1) + d_1T/2.3. \quad (33)$$

TABLE I. Values of parameters in viscosity-time equation.

Class designation	Temperature °C	$K_{1\alpha_1}K_{2\alpha_2}$ / $K_{1\alpha_1}+K_{2\alpha_2}$ poises^{-1}	$C(K_{1\alpha_1}g_1 - K_{2\alpha_2}g_2)$ / $f(g_1+g_2)$ poises^{-1}	$(K_{1\alpha_1}-K_{2\alpha_2})g_1g_2$ / g_1+g_2 minutes^{-1}
H. R. Lillie	447.8	1.26×10^{-14}	3.10×10^{-14}	2.41×10^{-2}
	505.5	3.80×10^{-13}	31.4×10^{-13}	3.27×10^{-3}
Dale and Stanworth	C_9 595	5.623×10^{-13}	193.9×10^{-13}	2.04×10^{-2}
	C_9 610	1.995×10^{-13}	54.2×10^{-13}	2.46×10^{-2}
	C_{40} 550	3.122×10^{-14}	398.1×10^{-14}	9.60×10^{-3}
	C_{40} 585	3.162×10^{-13}	174.6×10^{-13}	1.24×10^{-2}

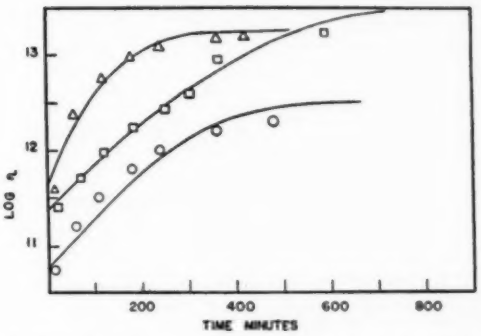


FIG. 5. Comparison of theory and experiment, log viscosity *vs.* time, for glass at different temperatures (after Dale and Stanworth, see reference 4). △ temperature 505°C, Glass No. *C* 40; □ temperature 550°C, Glass No. *C* 40; ○ temperature 595°C, Glass No. *C* 9.

This is the form of the equation which should be used when a calculation of parameters shows that ΔF_2^* is much greater than ΔF_1^*. With ordinary commercial glasses, this condition will hold when the temperature is in excess of approximately 1000°C.

An examination of the reduced viscosity equation, Eq. (33), reveals that it contains two constants, b_1 and ΔS_1^*, which are not separable. While a separation is desirable from the standpoint of an elucidation of the structure of glass, it is fortunately unnecessary if viscosities only are to be calculated. The same condition applies to the constants b_2 and ΔS_2^* in the general equation [Eq. (31)]. It is probable that this difficulty will arise, regardless of the manner in which the general viscosity equation [Eq. (19)] is expanded. For example, another expansion of that equation can be made in this manner:

$$\eta = h[(\exp(\Delta F_1^*/RT)/v_1) + (\exp(\Delta F_2^*/RT)/v_2)]$$

which can be written

$$\eta = e^{ax} + e^{by}$$

which is equivalent to

$$\eta = 2e^{(ax+by)/2}\cosh[(ax-by)/2].$$

We can then write

$$\eta = 2(h/b_1b_2)^{\frac{1}{2}}\exp\tfrac{1}{2}[(\Delta F_1^* + \Delta F_2^*)/RT + d_1T + d_2T] \times\cosh\tfrac{1}{2}[(\Delta F_1^* + \Delta F_2^*)/RT + \ln b_1/b_2 + d_1T - d_2T]$$

from which it can readily be seen that the same combinations of constants will appear.

The final assumption necessary to relate composition to viscosity is concerned with the manner in which each constituent oxide of the glass contributes to the value of each of the four parameters in the viscosity equation. It was found that the contribution of each oxide to the parameters with subscript 1 in Eqs. (31) and (33) was linearly proportional to the mole fraction of that oxide present. This is also true of the parameters with subscript 2 when alumina is not present. When it is present, it is necessary to assume that a compound with the

TABLE II. Composition of glasses used for the calculation of viscosity.

Glass designation Owen-Illinois	Composition mole percent										
	SiO_2	Al_2O_3	Fe_2O_3	B_2O_3	Na_2O	K_2O	CaO	MgO	SrO	SO_3	BaO
No. 17-1	0.7738	0.0018			0.1750				0.0478	0.0016	
17-2	0.7499	0.0006			0.1747				0.0731	0.0016	
17-3	0.7187	0.0013			0.1809				0.0975	0.0016	
17-4	0.7775	0.0123			0.1725				0.0362	0.0015	
17-5	0.7465	0.0125			0.1775				0.0620	0.0016	
17-6	0.7197	0.0133			0.1805				0.0849	0.0016	
17-7	0.7773	0.0244			0.1727				0.0240	0.0016	
17-8	0.7478	0.0249			0.1765				0.0491	0.0016	
17-9	0.7205	0.0254			0.1795				0.0230	0.0016	
15.5-1	0.7918	0.0013			0.1564				0.0490	0.0016	
15.5-2	0.7614	0.0006			0.1623				0.0749	0.0008	
15.5-3	0.7328	0.0006			0.1650				0.0999	0.0016	
15.5-4	0.7902	0.0124			0.1592				0.0367	0.0016	
15.5-5	0.7648	0.0131			0.1592				0.0621	0.0015	
15.5-6	0.7339	0.0127			0.1642				0.0876	0.0016	
15.5-7	0.7700	0.0253			0.1574				0.0372	0.0010	
15.5-8	0.7497	0.0258			0.1606				0.0629	0.0010	
15.5-9	0.7205	0.0263			0.1634				0.0887	0.0010	
14-1											
14-2	0.7793	0.0018			0.1458				0.0741	0.0010	
14-3	0.7497	0.0013			0.1474				0.1000	0.0016	
14-4	0.8092	0.0129			0.1402				0.0368	0.0008	
14-5											
14-6	0.7512	0.0127			0.1475				0.0869	0.0016	
14-7	0.7926	0.0247			0.1446				0.0372	0.0008	
14-8	0.7637	0.0258			0.1470				0.0619	0.0016	
14-9	0.7352	0.0261			0.1490				0.0872	0.0025	
10-*B*-0	0.7227	0.0006			0.1507		0.0643	0.0616			
10-*B*-1*S*	0.7149	0.0017		0.0093	0.1517		0.0622	0.0602			
10-*B*-3*S*	0.6866	0.0006		0.0265	0.1567		0.0647	0.0640			
10-*B*-5*S*	0.6808	0.0012		0.0428	0.1519		0.0627	0.0606			
10-*B*-1*D*	0.7293	0.0006		0.0086	0.1534		0.0551	0.0531			
10-*B*-3*D*	0.7367	0.0006		0.0232	0.1526		0.0438	0.0431			
10-*B*-5*D*	0.7405	0.0006		0.0425	0.1538		0.0312	0.0314			
10-*B*-1*N*	0.7248	0.0012		0.0076	0.1418		0.0631	0.0615			
10-*B*-3*N*	0.7317	0.0012		0.0230	0.1204		0.0634	0.0604			
10-*B*-5*N*	0.7305	0.0012		0.0401	0.1026		0.0636	0.0620			
12-*B*-0	0.7210	0.0012		0.0000	0.1319		0.0745	0.0715			
12-*B*-1*S*	0.7113	0.0012		0.0093	0.1331		0.0735	0.0716			
12-*B*-3*S*	0.6950	0.0012		0.0254	0.1322		0.0746	0.0716			
12-*B*-5*S*	0.6815	0.0006		0.0399	0.1305		0.0758	0.0718			
12-*B*-1*D*	0.7243	0.0012		0.0093	0.1322		0.0673	0.0658			
12-*B*-3*D*	0.7279	0.0012		0.0248	0.1343		0.0573	0.0546			
12-*B*-5*D*	0.7351	0.0012		0.0422	0.1345		0.0439	0.0431			
12-*B*-1*N*	0.7228	0.0012		0.0084	0.1221		0.0743	0.0713			
12-*B*-3*N*	0.7273	0.0012		0.0237	0.1027		0.0736	0.0716			
12-*B*-5*N*	0.7304	0.0012		0.0399	0.0830		0.0738	0.0718			
10-0	0.7227	0.0006			0.1507		0.0643	0.0616			
10-1*S*	0.7200	0.0012	0.0038		0.1518		0.0643	0.0606			
10-3*S*	0.7095	0.0012	0.0110		0.1536		0.0634	0.0613			
10-5*S*	0.6961	0.0012	0.0187		0.1563		0.0641	0.0635			
10-1*D*	0.7309	0.0012	0.0038		0.1530		0.0564	0.0548			
10-3*D*	0.7439	0.0012	0.0111		0.1557		0.0444	0.0437			
10-5*D*	0.7572	0.0012	0.0190		0.1585		0.0319	0.0322			
10-1*N*	0.7280	0.0012	0.0038		0.1429		0.0637	0.0605			
10-3*N*	0.7356	0.0012	0.0113		0.1260		0.0632	0.0626			
10-5*N*	0.7456	0.0012	0.0183		0.1089		0.0640	0.0619			

TABLE II.—*Continued.*

Glass designation Owen-Illinois	Composition mole percent										
	SiO_2	Al_2O_3	Fe_2O_3	B_2O_3	Na_2O	K_2O	CaO	MgO	SrO	SO_3	BaO
No. 12-0	0.7210	0.0012			0.1319		0.0745	0.0715			
12-1*S*	0.7163	0.0012	0.0037		0.1329		0.0740	0.0720			
12-3*S*	0.7038	0.0012	0.0109		0.1354		0.0759	0.0729			
12-5*S*	0.6915	0.0012	0.0190		0.1367		0.0766	0.0751			
12-1*D*	0.7261	0.0012	0.0037		0.1330		0.0698	0.0662			
12-3*D*	0.7389	0.0012	0.0113		0.1352		0.0581	0.0553			
12-5*D*	0.7512	0.0012	0.0192		0.1384		0.0459	0.0441			
12-1*N*	0.7163	0.0012	0.0037		0.1322		0.0741	0.0726			
12-3*N*	0.7312	0.0012	0.0112		0.1059		0.0766	0.0740			
12-5*N*	0.7409	0.0012	0.0189		0.0877		0.0764	0.0749			
Ba-0	0.7199	0.0012			0.1317		0.0744	0.0714		0.0015	0.0000
Ba-1*S*	0.7153	0.0012			0.1306		0.0759	0.0718		0.0015	0.0038
Ba-3*S*	0.7034	0.0012			0.1340		0.0756	0.0726		0.0015	0.0117
Ba-5*S*	0.6960	0.0006			0.1347		0.0755	0.0720		0.0015	0.0197
Ba-1*D*	0.7266	0.0006			0.1328		0.0686	0.0661		0.0015	0.0039
Ba-3*D*	0.7403	0.0012			0.1343		0.0570	0.0539		0.0015	0.0118
Ba-5*D*	0.7499	0.0012			0.1374		0.0459	0.0441		0.0015	0.0200
Ba-1*N*	0.7227	0.0012			0.1221		0.0749	0.0733		0.0015	0.0043
Ba-3*N*	0.7306	0.0012			0.1069		0.0756	0.0725		0.0015	0.0117
Ba-5*N*	0.7389	0.0012			0.0867		0.0775	0.0749		0.0015	0.0193
J. P. Poole											
1	0.7523				0.2477						
2	0.7601				0.1993		0.0405				
3	0.7594				0.1446		0.0960				
4	0.7568				0.0929		0.1503				
5	0.7074				0.2407		0.0519				
6	0.6629				0.2344		0.1028				
7	0.6135				0.2343		0.1522				
8	0.6979				0.2073		0.0948				
9	0.6514				0.1989		0.1496				
10	0.6126				0.1905		0.1969				
11	0.7356				0.2002			0.0462			
12	0.7724				0.1417			0.0860			
13	0.7720				0.1024			0.1255			
14	0.6969				0.2551			0.0483			
15	0.6747				0.2394			0.0859			
16	0.6035				0.2593			0.1372			
17	0.6962				0.2168			0.0870			
18	0.6481				0.2238			0.1281			
19	0.6223				0.2116			0.1661			
20	0.7455	0.0309			0.2236						
21	0.7385	0.0605			0.2010						
22	0.7214	0.0334			0.2452						
23	0.6833	0.0632			0.2535						
24	0.7352	0.0262			0.1361		0.1025				
25	0.7265	0.0559			0.1063		0.1112				
26	0.7374	0.0351			0.1507		0.0769				
27	0.7346	0.0648			0.1503		0.0503				
28	0.7122	0.0369			0.1482		0.1027				
30	0.7095	0.0377			0.2009		0.0519				
31	0.6572	0.0350			0.1996		0.1081				
32	0.6117	0.0351			0.2492		0.1040				
33	0.7572	0.0345			0.1522		0.0561				
34	0.6566	0.0364			0.1517		0.1553				
35	0.7500				0.1528		0.0521	0.0450			
36	0.6488				0.2533		0.0523	0.0458			
37	0.6640	0.0652			0.1555		0.1053				
38	0.7149	0.0207			0.1556		0.1088				
39	0.6699				0.2352				0.0950		
40	0.6346				0.2161				0.1493		
41	0.5988				0.2056				0.1956		
42	0.6771				0.2222						0.1006
43	0.6684				0.1794						0.1521
44	0.6027				0.2053						0.1920
45	0.7577				0.1233	0.1190					
46	0.7683					0.2317					

TABLE III. Comparison of observed and calculated viscosities.

Glass No. Owen-Illinois	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.
17-1	1727	2.000	1.960	1141	5.000	4.942	935	7.650	7.712
17-2	1683	2.000	1.779	1116	5.000	5.026	927	7.650	7.796
17-3	1605	2.000	2.046	1097	5.000	5.104	922	7.650	7.751
17-4	1822	2.000	1.965	1144	5.000	5.135	945	7.650	7.769
17-5	1700	2.000	2.101	1133	5.000	5.106	941	7.650	7.628
17-6	1655	2.000	2.115	1122	5.000	5.105	937	7.650	7.713
17-7	1833	2.000	2.086	1186	5.000	5.012	960	7.650	7.768
17-8	1752	2.000	2.206	1147	5.000	5.222	951	7.650	7.850
17-9	1689	2.000	2.250	1127	5.000	5.299	944	7.650	7.910
15-1	1797	2.000	1.907	1161	5.000	4.850	950	7.650	7.810
15-2	1700	2.000	1.966	1130	5.000	4.977	940	7.650	7.670
15-3	1650	2.000	1.971	1119	5.000	4.952	936	7.650	7.650
15-4	1850	2.000	2.003	1169	5.000	4.984	957	7.650	7.700
15-5	1761	2.000	2.091	1152	5.000	5.085	951	7.650	7.835
15-6	1661	2.000	2.179	1133	5.000	5.067	947	7.650	7.712
15-7	1900	2.000	2.107	1243	5.000	4.615	966	7.65	7.908
15-8	1777	2.000	2.203	1169	5.000	5.077	962	7.65	7.808
15-9	1689	2.000	2.292	1152	5.000	5.106	957	7.65	7.526
14-2	1761	2.000	1.941	1158	5.000	4.834	957	7.65	7.533
14-3	1683	2.000	2.004	1143	5.000	4.843	954	7.65	7.491
14-4	1858	2.000	2.131	1211	5.000	4.755	976	7.65	7.593
14-6	1745	2.000	2.056	1158	5.000	4.922	963	7.65	7.663
14-7	1939	2.000	2.137	1211	5.000	4.909	986	7.65	7.629
14-8	1827	2.000	2.194	1191	5.000	4.967	981	7.65	7.731
14-9	1755	2.000	2.211	1169	5.000	5.081	975	7.65	7.688
10-0	1747	2.000	2.008	1163	5.000	5.232	978	7.65	7.704
10-1*S*	1744	2.000	1.985	1155	5.000	5.232	975	7.65	7.696
10-3*S*	1700	2.000	2.074	1147	5.000	5.232	966	7.65	7.812
10-5*S*	1666	2.000	2.128	1139	5.000	5.226	961	7.65	7.804
10-1*D*	1750	2.000	1.984	1161	5.000	5.154	975	7.65	7.645
10-3*D*	1750	2.000	1.977	1161	5.000	5.042	964	7.65	7.734
10-5*D*	1783	2.000	1.881	1158	5.000	4.948	955	7.65	7.697
10-1*N*	1755	2.000	2.008	1172	5.000	5.064	983	7.65	7.666
10-3*N*	1777	2.000	2.017	1202	5.000	4.894	1000	7.65	7.495
10-5*N*	1852	2.000	1.956	1216	5.000	4.807	1014	7.65	7.400
12-0	1736	2.000	2.047	1194	5.000	5.054	1000	7.65	7.594
12-1*S*	1756	2.000	2.047	1183	5.000	5.112	994	7.65	7.678
12-3*S*	1705	2.000	2.139	1172	5.000	5.139	989	7.65	7.642
12-5*S*	1689	2.000	2.139	1164	5.000	5.142	983	7.65	7.734
12-1*D*	1755	2.000	2.047	1189	5.000	5.043	994	7.65	7.735
12-3*D*	1777	2.000	1.980	1189	5.000	4.938	986	7.65	7.606
12-5*D*	1794	2.000	1.929	1183	5.000	4.872	977	7.65	7.576
12-1*N*	1772	2.000	1.981	1191	5.000	5.049	1008	7.65	7.384
12-3*N*	1794	2.000	2.052	1222	5.000	4.869	1027	7.65	7.557
12-5*N*	1855	2.000	1.977	1244	5.000	4.629	1047	7.65	7.167
10-*B*-0	1747	2.000	2.008	1164	5.000	5.200	978	7.65	7.705
10-*B*-1*S*	1686	2.000	2.107	1152	5.000	5.252	974	7.65	7.742
10-*B*-3*S*	1636	2.000	2.072	1141	5.000	5.094	968	7.65	7.735
10-*B*-5*S*	1594	2.000	2.094	1116	5.000	5.118	963	7.65	7.970
10-*B*-1*D*	1722	2.000	2.021	1164	5.000	5.067	978	7.65	7.554
10-*B*-3*D*	1683	2.000	2.046	1144	5.000	5.055	973	7.65	7.563
10-*B*-5*D*	1666	2.000	1.919	1136	5.000	4.813	974	7.65	7.338
10-*B*-1*N*	1741	2.000	1.995	1169	5.000	5.120	981	7.65	7.751
10-*B*-3*N*	1739	2.000	1.994	1186	5.000	4.874	1002	7.65	7.585
10-*B*-5*N*	1755	2.000	1.890	1200	5.000	4.645	1014	7.65	7.567
12-*B*-0	1736	2.000	2.116	1194	5.000	5.057	1000	7.65	7.593

TABLE III.—*Continued.*

Glass No. Owen-Illinois	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.	Temp. °K	$\log\eta$ obs.	$\log\eta$ calc.
12-*B*-1*S*	1708	2.000	1.952	1177	5.000	5.989	994	7.65	7.593
12-*B*-3*S*	1666	2.000	2.092	1147	5.000	5.197	983	7.65	7.942
12-*B*-5*S*	1622	2.000	2.136	1133	5.000	5.168	979	7.65	8.057
12-*B*-1*D*	1750	2.000	2.001	1183	5.000	5.020	996	7.65	7.586
12-*B*-3*D*	1730	2.000	1.948	1161	5.000	4.999	991	7.65	7.540
12-*B*-5*D*	1716	2.000	1.839	1152	5.000	4.802	990	7.65	7.409
12-*B*-1*N*	1752	2.000	2.047	1186	5.000	5.678	1003	7.65	7.649
12-*B*-3*N*	1786	2.000	1.926	1180	5.000	4.969	1019	7.65	7.565
Ba-0	1737	2.000	2.110	1195	5.000	5.038	1001	7.65	7.535
Ba-1*S*	1738	2.000	2.077	1195	5.000	5.113	991	7.65	7.753
Ba-3*S*	1688	2.000	2.158	1168	5.000	5.180	985	7.65	7.757
Ba-5*S*	1698	2.000	2.067	1154	5.000	5.254	977	7.65	7.842
Ba-1*D*	1746	2.000	2.080	1188	5.000	5.069	994	7.65	7.580
Ba-3*D*	1753	2.000	2.000	1192	5.000	4.915	991	7.65	7.459
Ba-5*D*	1756	2.000	1.952	1183	5.000	4.895	975	7.65	7.612
Ba-1*N*	1739	2.000	2.119	1202	5.000	4.995	1008	7.65	7.491
Ba-3*N*	1770	2.000	2.064	1212	5.000	4.941	1018	7.65	7.436
Ba-5*N*	1842	2.000	1.928	1245	5.000	4.737	1050	7.65	7.009
J. P. Poole									
1	824	9.808	9.537	761	12.075	12.038			
2	850	10.077	10.107	783	12.578	12.790			
3	890	10.050	10.349	823	13.013	12.962			
4	948	10.124	9.857	851	13.798	13.494			
5	851	9.477	9.601	779	12.160	12.446			
6	845	9.820	9.936	782	12.788	12.490			
7	826	10.816	11.079	784	12.979	12.857			
8	848	10.354	11.014	784	13.357	13.040			
9	889	9.198	9.530	821	12.321	12.194			
10	889	9.417	10.069	822	12.753	12.727			
11	855	9.414	9.705	767	12.262	13.163			
16	855	10.659	11.227	767	13.910	14.797			
13	948	10.483	9.273	855	13.423	12.276			
14	805	10.183	10.177	754	12.151	12.258			
15	805	10.135	10.642	743	12.300	12.712			
16	806	10.211	10.300	753	12.441	12.339			
17	834	9.673	10.153	752	12.934	13.421			
18	854	9.462	9.414	766	13.014	12.667			
19	854	10.040	9.818	774	13.396	12.715			
20	853	9.768	9.959	768	12.879	13.320			
21	874	10.267	10.922	792	13.193	14.048			
22	853	9.565	9.458	765	12.822	12.889			
23	852	10.263	10.240	764	13.676	13.662			
24	953	9.645	9.435	854	13.218	12.934			
26	916	9.929	10.461	833	12.852	13.534			
28	915	10.052	10.830	853	12.421	13.106			
30	876	9.865	10.433	804	12.587	13.191			
31	875	10.302	10.877	827	12.278	12.716			
32	849	10.234	10.597	791	12.969	12.906			
33	915	9.788	10.277	822	12.917	13.732			
34	951	9.303	9.910	847	13.456	13.702			
35	873	10.152	10.592	800	12.672	13.925			
36	824	10.104	9.956	790	11.520	11.272			
38	916	9.719	10.017	832	12.989	13.161			
39	829	9.860	9.760	749	13.434	13.281			
40	829	10.026	10.271	766	13.060	13.116			
41	829	10.139	10.576	766	13.409	13.523			
42	795	10.148	10.329	717	13.960	14.108			
43	795	10.078	10.940	712	14.230	15.233			
44	794	9.954	10.127	714	13.952	14.047			

TABLE IV. Factors for calculation of parameters from chemical composition.

Constituent	ΔH_1	$\frac{\Delta S_1}{2.3R}+\log b_1$	$d_1/2.3$	ΔH_2	$\frac{\Delta S_2}{2.3R}+\log b_2$	$d_2/2.3$
SiO_2	77,000	−13.0	0.00395	164,000	7.26	0.00714
Al_2O_3	76,000	−30.0	0.00255	—	—	—
Fe_2O_3	800,000	207.0	0.00780	10,000	30.00	0.05400
B_2O_3	25,000	−11.0	0.00300	150,000	−17.70	−0.01500
$3Na_2O.Al_2O_3$	—	—	—	40,000	35.50	0.04200
CaO	37,000	−34.0	−0.00640	140,000	−29.10	−0.01900
MgO	37,000	−34.0	0.00640	80,000	−24.00	−0.00100
BaO	10,000	−30.0	−0.00600	136,000	1.00	−0.00400
SrO	36,000	−23.4	−0.00194	140,000	−29.20	−0.02800
Na_2O	30,000	−26.0	−0.00400	34,000	−8.72	−0.00090
SO_3	25,000	−20.0	0.00160	—	—	—

formula $3Na_2O \cdot Al_2O_3$ is formed in order to evaluate the parameters with subscript 2, and to recalculate mole fractions after making adjustment in composition which this assumption requires. The proportionality constants for the oxides are listed in Table IV. The product of this constant and the mole fraction of the corresponding oxide present will give the contribution of each oxide to each parameter.

The viscosities of a large number of glasses have been calculated by means of the above equations. The results of these calculations, together with the compositions of the glasses, expressed as mole percent, are shown in Tables II and III.

VII. DISCUSSION OF RESULTS

The agreement between observed and calculated logarithms of viscosity (Table III) is, in general, good. There are several possible reasons for the deviations which occur. An error of 0.1 percent in the chemical analysis for silica will introduce an error of almost three percent in the calculated logarithm of viscosity. It is unfortunate that this condition exists but it is probably impossible to obtain a viscosity-temperature equation which does not contain several terms, dependent upon temperature, that combine in such a way that analysis errors are greatly magnified. This condition, therefore, points to the need for extreme care in analysis if viscosities are to be calculated and indicates that analyses cannot be calculated from batch weights, as is so frequently done, if accuracy in predicting viscosities is desired.

There are other possible sources of error: it may be that the forms of the temperature dependent functions of the parameters in the viscosity equation are not correct. As has been pointed out, these forms are purely empirical and were used only because they yielded reasonable results. Further, only in the case of soda and alumina was any assumption made regarding the formation of compounds; other compounds may exist which exert an influence on the values of the parameters. It cannot be emphasized too strongly that the factors for the oxides, herein presented, are tentative values, based on the best available data on viscosity, but subject to change when better data become available or our knowledge of glass structure increases. A third possible source of error is the tacit assumption that no phase changes occur in the temperature ranges considered. If such changes do occur, it will be necessary to apply corrections to the entropy terms. It is also possible that a separation of the entropy and volume terms, ΔS and b, might allow a better calculation of viscosity values.

It should also be pointed out that by reason of the form of the model used to derive the various equations, a one-to-one ratio of units designated by the subscript 1 and the subscript 2, appears to exist. It is unlikely that this is so. The determination of the correct ratio, however, involves the derivation of the viscosity-time-temperature equations on a purely statistical basis. When work on such a derivation can be completed, it may serve to explain the discrepancies between the values of ΔF_1^*, found in this paper and the values found from breaking strength data and by other methods.

VIII. ACKNOWLEDGMENTS

Many individuals have contributed to this work. The authors wish to thank Dean H. Eyring and the staff of the Naval Research Project at the University of Utah for advice and criticism. Special appreciation is due Dr. Dan McLachlan, Department of Metallurgy, University of Utah, for his advice and encouragement and to Dr. O. G. Burch, Dr. E. F. Poncelet, Mr. Stephen Barber and others of the research staff of the Owens-Illinois Glass Company. The authors are indebted to Messrs. F. J. Bjorklund and O. L. Anderson, Department of Physics, University of Utah, for their assistance in the laborious task of calculating and checking factors for the oxides, and to Mr. James P. Poole, Pennsylvania State College, for supplying invaluable data on viscosity.

Circuit Analysis of Linear Varying-Parameter Networks

LOTFI A. ZADEH
Columbia University, New York, New York
(Received May 8, 1950)

This paper describes a theory of linear varying-parameter networks which is essentially a generalization of the familiar frequency domain theory of fixed linear networks. Such basic concepts as impedance, admittance, gain, etc., are extended to linear varying-parameter networks and their important properties are outlined. Extensions are given also of the general mesh and node equations, Thevenin's theorem, dualization, and some other relations that hold in the case of fixed networks. On the whole it is shown that many theorems, properties, and relations that hold in the case of fixed networks may be extended with proper modifications to linear varying-parameter networks.

I. INTRODUCTION

A LINEAR varying-parameter network is essentially a linear system in which one or more parameter-values are functions of time. Familiar examples of such networks are the various types of amplitude and phase modulators, variable bandwidth amplifiers, variable delay networks, etc. One may also add to this list the ordinary fixed linear networks, for these are simply a degenerate form of linear varying-parameter networks.

In mathematical terms, the behavior of a linear varying-parameter network is governed by a system of linear differential equations with time-dependent coefficients. These equations interrelate the voltages and currents in various parts of the network. A typical relation of this kind involving only two variables $u(t)$ and $v(t)$, each representing a voltage or a current, may be written in the following form:

$$[a_n(t)p^n+\cdots+a_1(t)p+a_0(t)]v(t) = [b_m(t)p^m+\cdots+b_1(t)p+b_0(t)]u(t), \quad (1)$$

where $p=d/dt$ and the a's and b's are known functions of time. The variable on the right, $u(t)$, is usually referred to as the input while the variable on the left, $v(t)$, is referred to as the output or the response to $u(t)$. The standard problem in connection with Eq. (1) is to find the response $v(t)$ to a given input $u(t)$ on the assumption that the network is unexcited prior to application of $u(t)$. The latter condition means, more precisely, that if $u(t)$ is zero for, say, $t<0$, then $v(t)$ must also be zero for $t<0$.

Most of the papers[1–7] having direct or indirect bearing on the subject of linear varying-parameter networks are concerned with various methods and techniques for obtaining approximate solutions to the problem stated above. In the present paper, however, a different aspect of the theory of linear varying-parameter networks will be considered. Specifically, the purpose is to extend to such networks the concepts, principles, procedures and, more generally, the techniques commonly used in the circuit analysis of fixed linear networks in the frequency domain. The basis for such an extension has been established in reference 6. Thus, by using the frequency analysis approach described in reference 6 it is possible to introduce into the theory of linear varying-parameter networks such familiar and useful notions as impedance, admittance, gain, phase, poles, zeros, etc.—all of these becoming time-varying quantities. It is also possible to extend to linear varying-parameter networks such well-known theorems as the superposition theorem, compensation theorem, Thevenin's theorem, etc., with practically no modifications. In addition it is possible to extend to linear varying-parameter networks the familiar and powerful techniques of dualization, impedance and frequency transformation,[8] spectrum analysis,[9] network expansion, vector representation in the complex plane, pole and zero analysis, etc.

The above list is by no means exhaustive, but is sufficient to indicate that many if not most of the concepts and techniques used in the circuit analysis of fixed networks may be extended to linear varying-parameter networks. Our objective in this paper is limited to that of extending only the more basic of such concepts and techniques. It appears that the exploration of the behavior and properties of linear varying-parameter networks from the frequency domain point of view offers a fruitful field for further study and investigations.

II. GENERAL THEORY

In order to simplify the terminology, the term "variable network" will frequently be used hereinafter in the same sense as "linear varying-parameter network."

As a preliminary to the extension of the basic frequency domain concepts such as impedance, admittance, etc., to linear varying-parameter networks, it will be necessary to introduce the notion of the so-called *system*

[1] J. R. Carson, Phys. Rev. **17**, 116 (1921).
[2] K. Dahr, *Die operatorenrechnung und ihre weiterentwicklung nebst anwendungen auf systeme variabler parameter* (Pettersons Bokindustri, Stockholm, 1938).
[3] L. A. Pipes, Phil. Mag. **25**, 585 (1938).
[4] C. Blanc, "Sur les equations differentielles lineaires a coefficients lentement variables," Bull. Tech. Suisse Romande **74**, 185 (1948).
[5] M. C. Gray and S. A. Schelkunoff, Bell Sys. Tech. J. **27**, 350 (1948).
[6] L. A. Zadeh, Proc. I.R.E. **38**, 291 (1950).
[7] L. A. Zadeh, J. App. Phys. **21**, 642 (1950).
[8] L. A. Zadeh, Proc. I.R.E. **38**, 1339 (1950).
[9] L. A. Zadeh, Proc. I.R.E. **38**, 1342 (1950).

function[10] of a variable network. Briefly, the system function of a variable network N is defined by the relation

$$H(j\omega;t)=\frac{v(t)}{u(t)}\Big]_{u(t)=e^{j\omega t}}. \tag{2}$$

In other words, the system function of N is a function $H(j\omega;t)$ such that $H(j\omega;t)e^{j\omega t}$ represents the response of N to $u(t)=e^{j\omega t}$. It is evident that for a fixed network $H(j\omega;t)$ reduces to $H(j\omega)$ and the above definition of $H(j\omega;t)$ reduces to the conventional definition of the system function of a fixed network.

The salient characteristic of the system function is that $H(j\omega;t)$ connects the output and input of a variable network in much the same manner as $H(j\omega)$ connects the output and input of a fixed network. Specifically, it may readily be shown that the output $v(t)$ of a variable network is given by the following relation:

$$v(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(j\omega;t)U(j\omega)e^{j\omega t}d\omega, \tag{3}$$

where $U(j\omega)$ is the Fourier transform of $u(t)$. It is seen that Eq. (3) is similar in form to the relation connecting the output and input of a fixed network:

$$v(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(j\omega)U(j\omega)e^{j\omega t}d\omega. \tag{4}$$

In view of this similarity and also of the fact that in Eq. (3) the variable t in $H(j\omega;t)$ plays the role of a parameter, it follows that the output of a variable network may be expressed in the same forms as the output of a fixed network. The more important of these forms are as follows:

(a) Fourier Transform Form

The Fourier transform equivalent of Eq. (3) is:

$$v(t)=\mathfrak{F}^{-1}\{H(j\omega;t)U(j\omega)\}, \tag{5}$$

where $\mathfrak{F}^{-1}$ represents the operation of inverse Fourier transformation, and the variable t in $H(j\omega;t)$ should be treated as if it were a constant parameter. It should be noted that while $v(t)$ is the inverse Fourier transform of $H(j\omega;t)U(j\omega)$, the latter quantity is not the Fourier transform of $v(t)$. In fact, it may readily be shown[6] that the Fourier transform of $v(t)$ is given by the relation

$$V(j\omega)=\int_{-\infty}^{\infty}\Gamma(j\omega';j\omega)U(j\omega')d\omega', \tag{6}$$

where the so-called *bi-frequency system function* $\Gamma(j\omega';j\omega)$ is the Fourier transform of $H(j\omega';t)e^{j\omega' t}$.

[10] The properties of this function are discussed in detail in reference 6.

(b) Laplace Transform Form

The Laplace transform equivalent of Eq. (3) is

$$v(t)=\mathfrak{L}^{-1}\{H(s;t)U(s)\} \tag{7}$$

where s is the complex frequency ($s=\sigma+j\omega$), $U(s)$ is the Laplace transform of $u(t)$, $\mathfrak{L}^{-1}$ represents the operation of inverse Laplace transformation, and the variable t in $H(s;t)$ should be treated as a constant parameter. For example, assume that the input is a unit impulse function (Dirac's δ-function) $\delta(t)$, i.e.,

$$u(t)=\delta(t), \tag{8}$$

and that $H(s;t)$ is of the form

$$H(s;t)=1/[s+\alpha(t)], \tag{9}$$

where $\alpha(t)$ is a given function of time. Since the Laplace transform of $\delta(t)$ is unity, it follows from Eq. (7) that

$$v(t)=\mathfrak{L}^{-1}\{1/[s+\alpha(t)]\} \tag{10}$$

and hence

$$v(t)=e^{-t\alpha(t)}, \tag{11}$$

which thus represents the response of the network to a unit impulse applied at $t=0$.

(c) Operational Form

The operational equivalent of Eq. (3) is:

$$v(t)=H(p;t)u(t), \tag{12}$$

where p is the usual differential operator. As in (a) and (b), in the process of operating by $H(p;t)$ on $u(t)$, the variable t in $H(p;t)$ should be treated as if it were a constant parameter.

From the above discussion it is seen that the system function of a variable network, particularly in its operational form, provides a very convenient means for describing the behavior of the network. Thus through use of the notion of the system function the behavior of a given variable network becomes identified with that of a linear time-dependent operator $H(p;t)$ which operates on the input to the network, $u(t)$, in accordance with the standard rules of operational calculus and thus yields the output $v(t)$. A significant advantage of the system function over other means for describing the behavior of a variable network is that once $H(j\omega;t)$ has been obtained, the response to a given input may be determined in most cases by using only a table of Laplace or Fourier transforms.

The problem of the determination of the system function of a variable network may be approached in several different ways. The most straightforward approach is to start with Eq. (1), which represents the relation between the output and input of the network. For convenience of the discussion, this relation will be rewritten in a compact form

$$L(p;t)v(t)=K(p;t)u(t), \tag{13}$$

where $L(p;t)$ and $K(p;t)$ stand respectively for the left-hand and right-hand operators in Eq. (1). Applying this relation to the case where $u(t)=e^{j\omega t}$, and making use of the fact that for this case the output is equal to $H(j\omega;t)e^{j\omega t}$, Eq. (13) gives,

$$L(p;t)\{H(j\omega;t)e^{j\omega t}\}=K(p;t)e^{j\omega t}. \tag{14}$$

By use of the identities

$$K(p;t)e^{j\omega t}\equiv e^{j\omega t}K(j\omega;t) \tag{15}$$

and

$$L(p;t)\{H(j\omega;t)e^{j\omega t}\}\equiv e^{j\omega t}L(p+j\omega;t)H(j\omega;t), \tag{16}$$

Eq. (14) reduces to the following relation:

$$L(p+j\omega;t)H(j\omega;t)=K(j\omega;t). \tag{17}$$

This relation is in effect a linear differential equation in which t is the independent variable, $j\omega$ is a fixed parameter, $K(j\omega;t)$ is the forcing function, and $H(j\omega;t)$ is the dependent variable. The problem of solving this equation for $H(j\omega;t)$ is considered in reference 6. Insofar as the purposes of the present analysis are concerned, it will be sufficient to assume that Eq. (17) can be solved for $H(j\omega;t)$ and thus yield an explicit expression for the system function of the network.

An indirect and frequently more convenient approach to the problem is provided by the techniques described in the present paper. In principle, the approach consists in expressing the system function of a given network in terms of the system functions of its constituent elements. This process is quite similar in form to the conventional circuit analysis of fixed networks and for this reason will be referred to by the same name. It should be emphasized, however, that the primary objective of this paper is not that of developing a method of obtaining an explicit expression for the system function of a variable network; rather, it is that of extending the conventional circuit analysis techniques to variable networks and, in particular, of establishing rules of combining the system functions of various basic combinations of networks.

An important rule of combination of system functions —one which will be found very useful later—concerns the manner in which successive operations by two or more system functions on a given input may be replaced by a single operation with a so-called *product* system function on the same input.

Consider two system functions $H_1(s;t)$ and $H_2(s;t)$ and let

$$v(t)=H_1(p;t)u(t) \tag{18}$$

and

$$w(t)=H_2(p;t)v(t). \tag{19}$$

These two relations may be combined into the one relation

$$w(t)=H_2(p;t)[H_1(p;t)u(t)], \tag{20}$$

which means that $w(t)$ is the result of successive operation by $H_1(p;t)$ and $H_2(p;t)$ on $u(t)$.

Now let it be assumed that Eq. (20) may be replaced by

$$w(t)=H_3(p;t)u(t) \tag{21}$$

where $H_3(p;t)$ is an as yet unknown system function such that the result of operation by $H_3(p;t)$ on $u(t)$ is identical with the result of successive operation by $H_1(p;t)$ and $H_2(p;t)$ on $u(t)$. $H_3(p;t)$ will be referred to as the *product system function* of $H_1(p;t)$ and $H_2(p;t)$.

To obtain the expression for $H_3(s;t)$ in terms of $H_1(s;t)$ and $H_2(s;t)$, let $u(t)=e^{st}$ in Eq. (20). Then from Eq. (20) the corresponding $w(t)$ is found to be

$$w(t)=H_2(p;t)[H_1(s;t)e^{st}] \tag{22}$$

or, equivalently

$$w(t)=e^{st}H_2(p+s;t)H_1(s;t). \tag{23}$$

On the other hand, substituting $u(t)=e^{st}$ into Eq. (21) we obtain

$$w(t)=H_3(s;t)e^{st}. \tag{24}$$

From comparison of Eqs. (23) and (24) it follows that the product system function of $H_1(s;t)$ and $H_2(s;t)$ is given by the operational relation

$$H_3(s;t)=H_2(p+s;t)H_1(s;t). \tag{25}$$

In this relation the variable t in $H_2(p+s;t)$ should be treated as a constant parameter and $H_1(s;t)$ should be treated as a function of time involving s as a parameter. It is evident that $H_2(p+s;t)$ should be interpreted as a *bilateral* operator, since in general $H_1(s;t)$ would not vanish for negative t.

Equation (25) may be written in a symbolic form

$$H_3=H_2*H_1 \tag{25a}$$

where the operation $*$ is distributive, associative and, in general, non-commutative. As should be expected, for the particular case of fixed networks Eq. (25) reduces to the familiar relation

$$H_3(s)=H_2(s)H_1(s). \tag{26}$$

The relation given by Eq. (25) expresses essentially the group property of system functions under multiplication. In connection with this property it is convenient to introduce at this point the notion of *inverse system functions*. Thus we shall say that two system functions $H_1(s;t)$ and $H_2(s;t)$ are the inverses of each other if and only if the associated product system function is equal to unity. The inverse of a system function $H(s;t)$ will be denoted as $H^{-1}(s;t)$ or $[H(s;t)]^{-1}$. With this notation the relation between a system function and its inverse assumes the following form:

$$H(p+s;t)H^{-1}(s;t)=H^{-1}(p+s;t)H(s;t)=1 \tag{27}$$

or, more simply,

$$H*H^{-1}=H^{-1}*H=1. \tag{27a}$$

This relation represents a generalization of the corre-

sponding relation in fixed networks, namely,

$$H(s)H^{-1}(s)=H^{-1}(s)H(s)=1. \tag{28}$$

Specific applications of the relations given above to the circuit analysis of linear varying-parameter networks will be made later. At this point it will suffice to note that the concept of the system function provides the necessary basis for the extension to variable networks of the frequency domain techniques commonly used in the analysis of fixed linear networks. In what follows we shall consider first such basic network functions as impedance, admittance, etc., and then will extend to variable networks some of the more important properties and relations that hold in the case of fixed networks.

III. BASIC NETWORK FUNCTIONS AND THEIR PROPERTIES

(a) Impedance

Consider a two-terminal variable network N. Let $v(t)$ be the voltage across the terminals of N and let $i(t)$ be the current flowing through N. The *instantaneous input impedance* of N or simply, the input impedance of N, is defined as a function $Z(j\omega;t)$ such that $Z(j\omega;t)e^{j\omega t}$ represents the output voltage $v(t)$ when the input current $i(t)$ is a complex exponential $i(t)=e^{j\omega t}$. In other words,

$$Z(j\omega;t)=\left.\frac{v(t)}{i(t)}\right]_{i(t)=e^{j\omega t}}. \tag{29}$$

It is seen that $Z(j\omega;t)$ is essentially the system function connecting $v(t)$ and $i(t)$. From this it follows immediately that the voltage $v(t)$ corresponding to a given input current $i(t)$ may be expressed in the four equivalent forms represented by Eqs. (3), (5), (7), and (12). For example, in operational form,

$$v(t)=Z(p;t)i(t), \tag{30}$$

where, as usual, the variable t in $Z(p;t)$ should be treated as if it were a constant parameter.

From the general relation represented by Eq. (17) it follows that if $v(t)$ and $i(t)$ are connected to each other by a differential equation such as Eq. (13), then $Z(s;t)$ satisfies the following differential equation

$$L(p+s;t)Z(s;t)=K(s;t). \tag{31}$$

A useful interpretation of $Z(j\omega;t)$ may be obtained by using the usual vector representation of voltages, currents, impedances, etc., in the complex plane. In such a plane $Z(j\omega;t)$ may be represented as a time-varying vector whose magnitude is $|Z(j\omega;t)|$ and whose argument is the phase of $Z(j\omega;t)$. This vector may be interpreted as the instantaneous amplitude of the output voltage $v(t)$ resulting from an input current of the form $i(t)=e^{j\omega t}$.

A very basic and quite obvious property of $Z(s;t)$ is related to the manner in which any number of impedances in series may be combined into a single impedance. Thus consider n variable networks $N_1, N_2, \cdots, N_n$ in series with each other, and let the input impedances of these networks be respectively $Z_1(s;t), Z_2(s;t), \cdots, Z_n(s;t)$. Because of the linear character of system functions it follows immediately that the input impedance of the over-all network is given by:

$$Z(s;t)=Z_1(s;t)+Z_2(s;t)+\cdots+Z_n(s;t). \tag{32}$$

As a simple illustration of this relation consider a network consisting of a series connection of a variable resistance $R=R(t)$, a variable inductance $L=L(t)$, and a variable capacitance $C=C(t)$. The impedances of these elements are respectively $R(t), L(t)s+\dot{L}(t)$, and $1/C(t)s$,[11] where the dot represents differentiation with respect to time. From Eq. (32) it follows that the input impedance in question is:

$$Z(s;t)=R(t)+L(t)s+\dot{L}(t)+1/[C(t)s]. \tag{33}$$

The input impedance of a variable network has many other properties in addition to those discussed above. Only a few of these will be considered in the following sections. The consideration of other properties is outside the scope of the present paper.

(b) Admittance

The notions of admittance and impedance of a two-terminal variable network are duals of each other. Thus, the input admittance of a network N is defined as a function $Y(j\omega;t)$ such that $Y(j\omega;t)e^{j\omega t}$ represents the current flowing through N when a voltage of the form $v(t)=e^{j\omega t}$ is applied across the input terminals of N. In other words,

$$Y(j\omega;t)=\left.\frac{i(t)}{v(t)}\right]_{v(t)=e^{j\omega t}}. \tag{34}$$

It is seen that $Y(j\omega;t)$ is the system function connecting $i(t)$ and $v(t)$. The dual of Eq. (30) is

$$i(t)=Y(p;t)v(t). \tag{35}$$

Also, the dual of Eq. (31) is

$$K(p+s;t)Y(s;t)=L(s;t). \tag{36}$$

Corresponding to the property of series addition of impedances there is the property of shunt addition of admittances. Thus, let $Y(s;t)$ be the admittance of a network consisting of a parallel connection of $N_1, N_2, \cdots, N_n$, the admittances of these networks being $Y_1(s;t), Y_2(s;t), \cdots, Y_n(s;t)$, respectively. It is evident that $Y(s;t)$ is given by

$$Y(s;t)=Y_1(s;t)+Y_2(s;t)+\cdots+Y_n(s;t) \tag{37}$$

which is the dual of Eq. (32).

An important mutual property of the input impedance

[11] These expressions follow from the relations $i=d(cv)/dt$ and $v=d(Li)/dt$.

and admittance of a variable network is that $Z(s;t)$ and $Y(s;t)$ are inverse system functions in the sense defined in section II. This fact follows immediately from the relations

$$v(t)=Z(p;t)i(t) \tag{30}$$

and

$$i(t)=Y(p;t)v(t). \tag{35}$$

Application of Eq. (27) to $Z(s;t)$ and $Y(s;t)$ yields the following relation between the impedance and admittance of a variable network

$$Z(p+s;t)Y(s;t)=Y(p+s;t)Z(s;t)=1 \tag{38}$$

or

$$Z*Y=Y*Z=1. \tag{39}$$

This relation represents a generalization of the familiar relation

$$Z(s)Y(s)=Y(s)Z(s)=1. \tag{39a}$$

Equation (38) may be used for determining the admittance of a network from the knowledge of its impedance and *vice-versa*. For example, consider a variable capacitor the capacitance of which is a function of time, $C=C(t)$. The impedance of such a capacitor is

$$Z(s;t)=1/[C(t)s]. \tag{40}$$

From Eq. (38) it follows that the admittance of this capacitor satisfies the operational relation

$$\frac{1}{C(t)(p+s)}\{Y(s;t)\}=1. \tag{41}$$

Multiplying both sides of Eq. (41) first by $C(t)$ and then by $(p+s)$ we obtain

$$Y(s;t)=(p+s)C(t) \tag{42}$$

or

$$Y(s;t)=C(t)s+\dot{C}(t), \tag{43}$$

which thus represents the admittance of a variable capacitor.

In general, the determination of the impedance of a variable network from knowledge of its admittance, or *vice-versa*, is not an easy problem. For the purposes of the present analysis it will suffice to note that in most cases the solution of this problem reduces to the solution of an integral equation which results from the application of Eq. (3) to Eq. (38). In many practical cases it is not difficult to obtain an adequate approximation to $Z(s;t)$ when $Y(s,t)$ is given, or *vice-versa*, either directly from Eq. (38) or by using an approximate method of solution of the integral equation referred to above.

(c) Gain (Transfer Function)

Let $e_1(t)$ and $e_2(t)$ denote respectively the input and output of a four-terminal variable network N. By analogy with the definitions of impedance and admittance, the *gain* or *transfer function* of N is defined as a function $G(j\omega;t)$ such that $G(j\omega;t)e^{j\omega t}$ represents the output voltage $e_2(t)$ when the input voltage $e_1(t)$ is of the form $e_1(t)=e^{j\omega t}$. In other words,

$$G(j\omega;t)=\frac{e_2(t)}{e_1(t)}\Big]_{e_1(t)=e^{j\omega t}}. \tag{44}$$

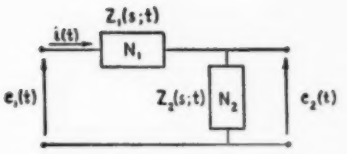


FIG. 1. Calculation of the gain of a variable network.

It is seen that $G(j\omega;t)$ is the system function connecting $e_2(t)$ and $e_1(t)$. As in the case of $Z(j\omega;t)$ and $Y(j\omega;t)$, we may write

$$e_2(t)=G(p;t)e_1(t). \tag{45}$$

Formally, the expression for the gain of a variable network may be obtained in much the same manner as in the case of a fixed network. Consider for example the network shown in Fig. 1 in which N_1 and N_2 are two variable networks whose impedances are respectively, $Z_1(s;t)$ and $Z_2(s;t)$; $e_1(t)$ and $e_2(t)$ represent respectively the input and output voltages, and the problem is to find an expression for the gain $G(s;t)$ in terms of the component impedances of the network.

The input admittance of the network under consideration may be written as:

$$Y(p;t)=[Z_1(p;t)+Z_2(p;t)]^{-1} \tag{46}$$

and hence the input current $i(t)$ may be expressed as

$$i(t)=[Z_1(p;t)+Z_2(p;t)]^{-1}e_1(t). \tag{47}$$

The output voltage $e_2(t)$ is related to $i(t)$ through

$$e_2(t)=Z_2(p;t)i(t). \tag{48}$$

By combining Eqs. (48) and (47), and making use of the multiplication rule of system functions (Eq. (25)) it follows that the expression for the gain is:

$$G(s;t)=Z_2(p+s;t)[Z_1(s;t)+Z_2(s;t)]^{-1} \tag{49}$$

or

$$G=Z_2*(Z_1+Z_2)^{-1}. \tag{49a}$$

The corresponding expression for a fixed network reads:

$$G(s)=Z_2(s)[Z_1(s)+Z_2(s)]^{-1}. \tag{50}$$

A network configuration frequently used in practice consists of two or more four-terminal networks which are connected in tandem. Consider two such networks N_1 and N_2 (Fig. 2) and let their gains be respectively $G_1(s;t)$ and $G_2(s;t)$. From Eq. (25) it follows immediately that the overall gain $G(s;t)$ is given by the relation

$$G(s;t)=G_2(p+s;t)G_1(s;t). \tag{51}$$

In this relation the variable t in $G_2(p+s;t)$ should be treated as if it were a constant parameter and $G_1(s;t)$

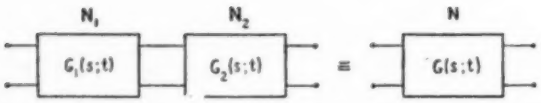


FIG. 2. Tandem connection of two variable networks.

should be treated as a function of time involving s as a parameter. It will be noticed that for the particular case of fixed networks Eq. (51) reduces to the familiar relation

$$G(s)=G_2(s)G_1(s). \tag{52}$$

As an application of Eq. (51) consider the case where N_1 is an amplitude modulator whose gain is

$$G_1(s;t)=A(t) \tag{53}$$

and N_2 is a phase modulator whose gain is expressed by

$$G_2(s;t)=e^{-s\alpha\cos\beta t} \tag{54}$$

where α and β are constants and $A(t)$ is a specified function of time.

From Eq. (51) the gain of the over-all system is found to be

$$G(s;t)=e^{-(p+s)\alpha\cos\beta t}\{A(t)\} \tag{55}$$

or

$$G(s;t)=e^{-s\alpha\cos\beta t}A(t-\alpha\cos\beta t). \tag{56}$$

It will be noted that in general an expression of the form

$$G(s;t)=G_2(s;t)G_1(t) \tag{57}$$

represents a composite network consisting of a variable network whose gain is $G_2(s;t)$ *followed* by an amplitude modulator whose gain is $G_1(t)$. Thus, Eq. (57) shows that an amplitude modulator *followed* by a phase modulator (the order of N_1 and N_2 in Fig. 2) is equivalent to the same phase modulator *followed* by an amplitude modulator whose gain is the same as that of the original modulator except for a delay in time. The same conclusion could be reached, of course, by purely physical considerations.

IV. BASIC NETWORK PROPERTIES

As was stated in the introduction, many if not most of the theorems, properties, relations, etc., that hold in the case of fixed networks, may be extended to linear varying-parameter networks. In what follows only a few such theorems, properties, relations, etc., will be considered. The topics discussed below have been chosen primarily because of their simplicity, basic importance and illustrative value.

(a) General Mesh and Node Equations of a Variable Network

The mesh and node equations of a variable network have the same form as those of a fixed network, the only difference being that all impedances and admittances in the former type of network are time-dependent quantities. Thus, for example, using operational notation the mesh equations of a variable network may be written in the following form:

$$\begin{aligned}
e_1&=Z_{11}(p;t)i_1+Z_{12}(p;t)i_2+\cdots+Z_{1n}(p;t)i_n\\
e_2&=Z_{21}(p;t)i_1+Z_{22}(p;t)i_2+\cdots+Z_{2n}(p;t)i_n\\
&\cdots\cdots\cdots\cdots\cdots\\
e_n&=Z_{n1}(p;t)i_1+Z_{n2}(p;t)i_2+\cdots+Z_{nn}(p;t)i_n.
\end{aligned} \tag{58}$$

In these equations $i_1, i_2, \cdots, i_n$ represent the mesh currents; $e_1, e_2, \cdots, e_n$ represent the loop voltages; and $Z_{11}(p;t)$, $Z_{12}(p;t)$, etc., represent the self and mutual loop impedances.

It is important to note that the principle of reciprocity does hold in the case of linear varying-parameter networks. Thus, in Eq. (58) $Z_{\nu\mu}(p;t)$ is equal to $Z_{\mu\nu}(p;t)$.

(b) Thevenin's Theorem

Thevenin's theorem may be extended to variable networks with practically no modifications in its standard form. Thus let aa' be the terminals of a variable network N (Fig. 3) and let $Z_L(s;t)$ be an impedance connected across aa'. *Then, insofar as $Z_L(s;t)$ is concerned, the network N may be replaced by a voltage source $e(t)$ in series with an impedance $Z(s;t)$, where $e(t)$ is the open-circuit voltage across aa' and $Z(s;t)$ is the impedance looking into aa' when all sources of voltage inside N are short-circuited.*

To prove the above statement we shall assume for simplicity that there is only one voltage source $e_s(t)$ inside N. Let $v(t)$ and $i(t)$ denote respectively the voltage across aa' and the current through these terminals. Because of the linearity of the system we can write

$$i(t)=Y(p;t)v(t)+Y'(p;t)e_s(t), \tag{59}$$

where $Y(p;t)$ and $Y'(p;t)$ are some as yet unspecified admittances. To obtain the physical significance of $Y(p;t)$ and $Y'(p;t)$ we proceed as follows:

First, let the voltage source $e_s(t)$ be short-circuited, i.e., $e_s(t)=0$. Then, Eq. (59) gives

$$i(t)=Y(p;t)v(t) \tag{60}$$

which shows that $Y(p;t)$ is the admittance looking into aa' when $e_s(t)$ is short-circuited.

Next, let the terminals aa' be open-circuited, i.e., $i(t)=0$, and let the voltage appearing across aa' under these conditions be denoted by $e(t)$. For this case Eq. (59) gives

$$0=Y(p;t)e(t)+Y'(p;t)e_s(t). \tag{61}$$

Combining Eqs. (59) and (61) we find

$$i(t)=Y(p;t)[v(t)-e(t)] \tag{62}$$

or, equivalently

$$v(t)-e(t)=Z(p;t)i(t) \tag{63}$$

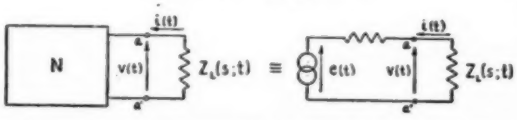


FIG. 3. Extension of Thevenin's theorem to variable networks.

where in view of Eq. (60) $Z(p;t)$ represents the impedance looking into aa' when all voltage sources in N are short-circuited. Equation (63) constitutes essentially a proof of the stated extension of Thevenin's theorem to linear varying-parameter networks.

(c) Dualization

A widely used technique in both the analysis and synthesis of fixed networks is that of dualization. This technique may also be used in variable networks in much the same manner as in the case of fixed networks. In fact, the only new aspect of dualization in the case of variable networks is that the relations between dual circuit-elements assume the following form:

Dual of a capacitance $C=f(t)$ is an inductance $L=f(t)$ and vice-versa.

Dual of resistance $R=h(t)$ is a conductance $G=h(t)$ and vice-versa.

It will be noticed that for the particular case of fixed networks the above relations reduce to the standard form. It should be emphasized that all dual properties and relations that apply to fixed networks, apply also to linear varying-parameter networks. For example, if the input impedance of a network N is $Z(p;t)$, then the input impedance of the dual of N is the inverse of $Z(p;t)$.

V. CONCLUDING REMARKS

The results obtained in the preceding sections show that the basic notion of the system function and the derived notions of impedance, admittance, gain, etc., provide the necessary basis for the extension of the conventional "fixed network" techniques of circuit analysis to linear varying-parameter networks. Much further work along these lines remains to be done, particularly in connection with the difficult problem of finding the inverse of a given system function. Nevertheless, even at its present stage of development, the circuit analysis approach outlined in this paper not only enhances understanding of the behavior of variable networks but may also be used with advantage in the solution of many problems arising in connection with the analysis and synthesis of linear varying-parameter systems.

Convection Currents in Porous Media. III. Extended Theory of the Critical Gradient

F. T. Rogers, Jr.
U. S. Naval Ordnance Test Station, Inyokern, California*

AND

H. L. Morrison
North Carolina State College of Engineering, Raleigh, North Carolina

(Received April 19, 1950)

The theory of the critical gradient for onset of thermal convection in a fluid entrapped in a porous medium is extended to allow for the exponential dependence of viscosity upon temperature and for non-linear vertical temperature-distributions which characterize transients in heat flow. Although there are approximations in the mathematical treatment, the extended theory agrees rather well with experimental data, as the simple theory does not in certain instances. New experimental data are reported for the critical gradient, obtained largely with silicone fluids in unconsolidated sands.

An earlier paper[1] under this same general title developed from basic considerations the expression

$$\beta_c = 4\pi^2 h^2 \mu / kg\rho_0 \alpha D^2, \qquad (1)$$

for the (negative) mean thermal gradient which must obtain in a fluid entrapped in a porous medium in order for convection to occur. Here h^2 represents the thermal diffusivity of fluid and medium *in situ*, k the flow-permeability of the medium, μ and ρ_0 the viscosity and bottom density of the fluid, α the coefficient of cubical thermal expansion of fluid relative to medium, and g the acceleration of gravity. In the development μ was for simplicity taken to be everywhere constant, and the temperature was taken to be essentially linear in the distance above the bottom of fluid and medium.

A subsequent paper[2] reported the results of experiments designed to test the validity of Eq. (1). It was clear from these experiments that Eq. (1) is substantially correct, except for its prediction of excessive[1] β_c-values, but that its applicability is severely limited by the underlying assumptions. In the experiments μ was of necessity not constant, and in most of them convection began before a linear (i.e., steady-state) condition of heat flow could be established. The viscosity depended upon temperature according to

$$\mu \cong \mu_0 \exp[-\nu(\Theta - \Theta_0)], \qquad (2)$$

* Post office: China Lake, California.

[1] C. W. Horton and F. T. Rogers, Jr., J. App. Phys. **16**, 367 (1945). A paper by E. R. Lapwood, Proc. Camb. Phil. Soc. **44**, 605 (1948), carried out this and related developments in a somewhat more comprehensive manner.

[2] H. L. Morrison, F. T. Rogers, Jr., and C. W. Horton, J. App. Phys. **20**, 1027 (1949).

where ν is a constant for a given fluid and μ_0 is the viscosity at temperature Θ_0, whereas data such as are shown in Fig. 1 imply departures of vertical temperature-distribution from linearity given by

$$\theta \cong \theta_1 \sin(\pi z/D)+\theta_2 \sin(2\pi z/D), \tag{3}$$

θ_1 and θ_2 being constants, z being distance above bottom, and D being bottom-to-top distance through the convective fluid. That the experimental results could be represented by the empirical equation,

$$\beta_{c,\,\mathrm{obs}} \cong \frac{4\pi^2 h^2 \mu_{\bar{h}}}{kg\rho_0 D^2}\left(\frac{k\rho_0}{h^2\mu_{\bar{h}}K}\right)^{0.51}, \tag{4}$$

where $K=10^{-3}$ sec.^{2}/cm^2 deg. C for c.g.s. units, bears out both points made at the beginning of this paragraph; here $\mu_{\bar{h}}$ is the viscosity corresponding to the average temperature in the convective fluid.

THE EXTENDED THEORY

The object of extending the simple theory for β_c is to obtain an improved form of Eq. (1), which accounts for experimentally observed critical gradients as well as does Eq. (4), and perhaps to find a physical basis for Eq. (4). The process of extension is merely to allow for the sense of Eqs. (2) and (3) in a repetition of the calculations in reference 1. Thus Eq. (14) of reference 1 now appears as

$$\frac{\partial^2 w}{\partial z^2}-\nu\beta\frac{\partial w}{\partial z}-Aw+A\lambda\theta\exp[\nu(\beta z+\theta)]=0, \tag{5}$$

where $A=l^2+m^2$ and $\lambda=kg\rho_0\alpha/\mu_0$, μ_0 being identified as the bottom viscosity. Likewise, Eq. (15) of reference 1 is now

$$n\theta+w\beta+w\frac{\partial\theta}{\partial z}=h^2\left(\frac{\partial^2}{\partial^2 z}-l^2-m^2\right)\theta=0, \tag{6}$$

reflecting the fact that θ-derivatives may be comparable to β itself. Since the precision of available data does not seem sufficient to warrant retention of the second term in Eq. (3), the sense of Eq. (3) is, approximately, that $\theta=\theta_m \sin sz$; use of this in Eq. (6) leads, for small w and for $n=0$ to

$$\theta=-\frac{w(\beta+s\theta_m)}{h^2\sigma}, \tag{7}$$

where $\sigma=s^2+l^2+m^2$. Then substitution of Eq. (7) into Eq. (5) yields ($\beta=-\beta_c$)

$$\frac{\partial^2 w}{\partial z^2}+\nu\beta_c\frac{\partial w}{\partial z}-Aw+\frac{A}{B}(\beta_c-s\theta_m)e^{-\nu\beta_c z}w=0, \tag{8}$$

where $B=h^2\sigma/\lambda$, which must be solved for β_c under the boundary conditions $w(0)=w(D)=0$.

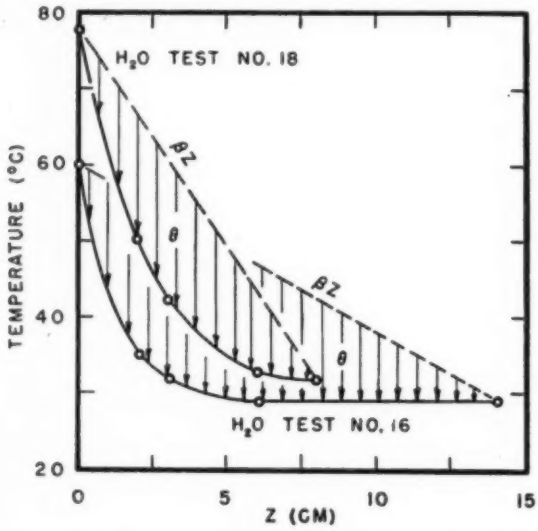


FIG. 1. Resolution of two typical non-linear vertical distributions of temperature, into linear and non-linear portions; see Eq. (3). These are the same distributions as are shown in Fig. 1 of reference 1. (Note: the axis of abscissae in Fig. 1 of reference 1 should have been labelled $D-z$ rather than z.)

Case I: $\nu=0$

If the viscosity were actually constant, as was assumed in reference 1, but if the vertical temperature distribution were not linear, then the first derivative and the exponential in Eq. (8) would be missing. In this event Eq. (8) would have as its solution a linear combination of sine and cosine with the argument $z(A/B)^{\frac{1}{2}}\times(\beta_c-s\theta_m-B)^{\frac{1}{2}}$, and in order that boundary conditions be satisfied, this argument would equal π for $z=D$. Hence, since $A=\pi/D$ and $\sigma=2A$ by arguments advanced in reference 1,

$$\beta_c=\frac{4\pi^2 h^2\mu_0}{kg\rho_0\alpha D^2}+\frac{\pi\theta_m}{D}, \quad \nu=0. \tag{9}$$

Equation (9) clearly shows how, as reference 2 reported, a negative non-linearity, θ_m, can lead to drastically smaller critical gradients than Eq. (1) would predict.

Case II: νBD small

For $\nu>0$ but νBD small, Eq. (8) has the series solution

$$w=\sum_{j=1}^{\infty} a_j z^j,$$

where

$$\left.\begin{aligned} a_j&=\left[\frac{Q^{\frac{1}{2}(j-1)}}{(j-2)!}-\cdots\right]\frac{a_1}{j(j-1)!}, \quad j \text{ odd},\\ a_j&=\left[\frac{Q^{\frac{1}{2}(j-2)}}{(j-2)!}-\cdots\right]\frac{\nu\beta a_1}{j(j-1)!}, \quad j \text{ even},\end{aligned}\right\}$$

and

$$Q=A+A(s\theta_m-\beta_c)/B,$$

TABLE I. New observations of the critical gradient, with $D=10$ cm and $k=140\times10^{-8}$ c.g.s. All units are c.g.s.

Liquid	Test No.	$10^3\times\alpha$	$10^3\times h^2$	ρ	z_0 (cm)	Temperatures at onset $z=0$	2	6	10	μ_{hr} (poises)	Observed β_c
Glycerol	5	0.495	2	1.26	3.6	100	66	44.5	38.5	0.62	6.15
Glycerol	6	0.495	2	1.26	3.6	90	60.2	36	34.8	0.98	5.52
1-DC-200	2	1.38	1.6	0.818	4.1	38.5	31	27.9	27.8	0.0074	0.93
1-DC-200	3	1.38	1.6	0.818	3.8	30	24.8	22.2	22	0.0080	0.80
1-DC-200	5	1.38	1.6	0.818	3.3	30	26	24	23.7	0.0080	0.63
10-DC-200	1	1.12	1.7	0.940	4.0	67	44	31	29.6	0.080	3.1
10-DC-200	3	1.12	1.7	0.940	4.1	70	47.6	34.2	32.8	0.078	3.72
10-DC-200	5	1.12	1.7	0.940	4.1	50	38	29	27.3	0.085	2.27
100-DC-200	1	0.923	1.7	0.968	3.0	90	56	34.4	32	0.67	5.80
100-DC-200	4	0.923	1.7	0.968	3.1	85	50	28	28	0.74	5.40
100-DC-200	6	0.923	1.7	0.968	3.3	100	60	34.5	33.6	0.675	6.60

to the first order in νBD. By suitable manipulation, this can be rearranged, so that the boundary condition $w(D)=0$ leads to

$$\pi[(\beta_c/B)-(s\theta_m/B)-1]^{\frac{1}{2}}\sin\pi[(\beta_c/B)-(s\theta_m/B)-1]^{\frac{1}{2}} -\nu\beta_c D\cos\pi[(\beta_c/B)-(s\theta_m/B)-1]^{\frac{1}{2}} +\text{higher-order terms in } \nu BD=0. \quad (10)$$

The first two terms in Eq. (10) suffice to determine β_c for the present case; since $\beta_c/B-s\theta_m/B$ must be close to 2, some reduction yields

$$\beta_c=\frac{\dfrac{4\pi^2h^2\mu_0}{kg\rho_0\alpha D^2}+\dfrac{\pi\theta_m}{D}}{1+\dfrac{2\nu D}{\pi^2}\dfrac{4\pi^2h^2\mu_0}{kg\rho_0\alpha D^2}}, \quad \nu BD \text{ small.} \quad (11)$$

The outstanding feature of Eq. (11) is its insistence that a non-constant viscosity should reduce the critical gradient.

Case III: νBD large

Equation (8) can be put into the form of Bessel's equation[3] by the transformations $\eta=w\exp(\nu\beta_c z)$ and $\zeta=[4A(\beta_c/B-s\theta_m/B)/\nu^2\beta_c^2]^{\frac{1}{2}}\exp(-\frac{1}{2}\nu\beta_c z)$,

$$\eta''+\frac{\eta'}{\zeta^2}+\left[1-\frac{1+(4A/\nu^2\beta_c^2)}{\zeta^2}\right]=0, \quad (12)$$

and this is well suited to the study of less simple convective environments. In terms of

$$q=(1+4A/\nu^2\beta_c^2)^{\frac{1}{2}},$$
$$\zeta_0=2\pi(\beta_c/B-s\theta_m/B)^{\frac{1}{2}}/\nu\beta_c D,$$
and
$$\zeta_D=\zeta_0\exp(-\tfrac{1}{2}\nu\beta_c D),$$

its solution, a linear combination of Bessel functions of

TABLE II. (Negative) critical gradients calculated by various forms of the theory.

Liquid	Test No.	μ_0 (poises)	$2B$ (°C/cm)	ν (°C^{-1})	θ_m (°C)	β_c, Critical gradients at onset (°C/cm) Observed	Calculated by Eq. (1)	Eq. (4)	Eq. (11)	Eq. (14)
Glycerol	2	0.0012	1.60	0.0860	40	9.4	550	14	—	6.5
Glycerol	5	0.0148	13.76	0.0860	25	6.15	572	11	4.76	(1.76)
Glycerol	6	0.0316	29.2	0.0860	27	5.52	904	14	13.75	(1.23)
CCl_4	3	0.0084	2.77	0.0126	0	0.8	2.86	1.1	2.58	(0)
Water	1	0.00448	16.34	0.0142	13	3.8	26.3	3.1	8.33	—
Water	5	0.00303	19.72	0.0142	20	6.1	34.1	3.5	8.57	—
Water	6	0.00385	19.94	0.0142	18	5.4	38.4	3.7	9.08	—
Water	7	0.00436	15.88	0.0142	14	3.95	26.3	3.1	7.88	—
Water	8	0.00436	15.88	0.0142	15	4.1	27.0	3.1	7.63	—
Water	9	0.00469	8.72	0.0142	17	2.2	14.1	1.6	3.62	—
Water	11	0.00317	18.06	0.0142	25	6.9	33.0	4.4	5.81	—
Water	13	0.0028	436	0.0142	—	>7.5	460	<12	<32	(>0?)
Water	15	0.00412	15.00	0.0142	17	3.75	24.3	3.0	6.75	—
Water	16	0.00468	8.72	0.0142	15	2.2	14.1	1.6	3.96	(6.25)
Water	18	0.00368	20.94	0.0142	20	5.6	40.2	4.7	8.82	(10.79)
1-DC-200	2	0.0072	2.94	0.0120	6.5	0.93	3.02	1.1	0.84	(13.88)
1-DC-200	3	0.0080	3.27	0.0120	4	0.80	3.26	1.1	1.84	—
1-DC-200	5	0.0080	3.26	0.0120	2.5	0.63	3.26	1.1	2.27	—
10-DC-200	1	0.0475	22.08	0.0152	23	3.1	36.4	4.7	(8.85)	7.50
10-DC-200	3	0.0469	21.80	0.0152	19	3.72	40.0	4.5	(9.46)	6.45
10-DC-200	5	0.0601	27.94	0.0152	9.5	2.27	39.1	4.5	(13.43)	4.30
100-DC-200	1	0.322	176.6	0.0177	26.5	5.80	367	16.4	(22.94)	2.12
100-DC-200	4	0.362	198.6	0.0177	30	5.40	407	17.1	(23.58)	2.46
100-DC-200	6	0.287	157.4	0.0177	33	6.60	370	17.6	(22.39)	2.90

[3] H. Massey and C. Mohr, Proc. Roy. Soc. London **A152**, 693 (1935).

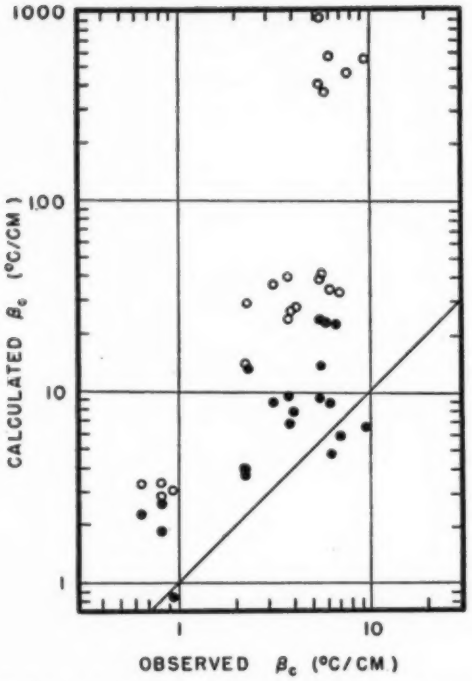


FIG. 2. Comparison of the simple theory, Eq. (1) (open circles), and the extended theory, Eqs. (11) and (14) (filled circles), with observational data. Points on the straight line would represent perfect agreement. The general tendency of points to lie above the line was anticipated in reference 1, and is attributable in part to a choice of boundary conditions differing somewhat from those which obtained in the experiments.

the first and second kinds of order q, can be written as

$$J_q(\zeta_0)J_{-q}(\zeta_D)-J_q(\zeta_D)J_{-q}(\zeta_0)=0, \qquad (13)$$

and in this form satisfies the condition $w(D)=0$. Since the domain of large $\nu\beta_c D$-values is not of great interest at present, it suffices to indicate what Eq. (13) predicts for the typical value $\nu\beta_c D=5.62$ (i.e., $q=\frac{3}{2}$); in this instance Eq. (13) reduces to

$$\zeta_0\cong 2^{\frac{1}{2}}\pi,$$

whence for θ_m sufficiently negative

$$\beta_c|_{\theta_m\ll 0}\cong\frac{1}{\nu D}\left(\frac{-4\pi\theta_m/D}{4\pi^2h^2\mu_0/kg\rho_0\alpha D^2}\right)^{\frac{1}{3}}, \quad \nu\beta_c D=5.62. \qquad (14)$$

Computational experience suggests that Eq. (14) holds as a solution of Eq. (13), in an approximate sense for $3<\nu\beta_c D<9$.

COMPARISON WITH EXPERIMENT

For the purpose of providing a larger body of data against which to compare the above extended theory, additional experiments have been carried out by the technique reported in reference 2. In most of these experiments silicone fluids were used, each being identified by type (Dow-Corning DC-200) and (by a suitable prefix) by its nominal kinematic viscosity in centistokes at 25°C. Ciba oil red C Y dye was employed[4] in the silicone experiments as a visible indicator of the state of fluid motion, and an unidentified blue dye in the few glycerol experiments. Table I exhibits typical data obtained from these new experiments.

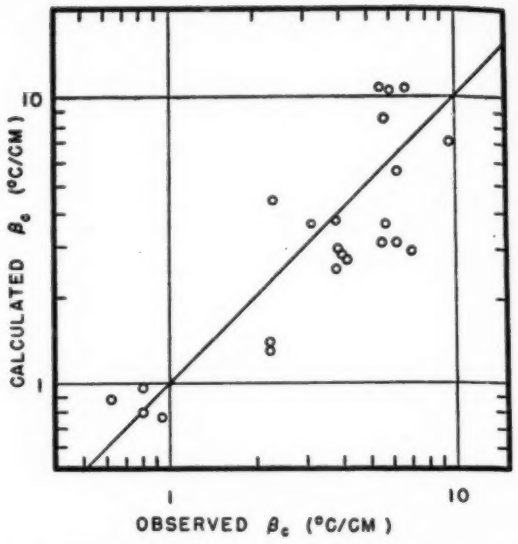


FIG. 3. Comparison of the improved empirical formula, Eq. (15), with observational data. Points on the straight line represent perfect agreement. Although Eq. (15) predicts critical gradients which depart in some instances by a factor of three from observed values, it is perhaps the best formula available at the present time.

Table II contains the observed mean critical gradients for all tests reported in reference 2 as well as for those in Table I, together with gradients calculated by the various formulas of theory.

It is obvious by inspection of Table II that the extended theory as embodied in Eqs. (11) and (14), is in far better accord with the observational data than is the original simple theory summarized by Eq. (1). Figure 2 illustrates this, and at the same time demonstrates the quasi-statistical character of the accordance. For the present it does not seem worth while to expend much effort on mathematical refinements of the extended theory, because many components of experimental technique deserve prior improvement.[5]

While the entries in Table II show that the empirical Eq. (4) is in about as good agreement with observation as are Eqs. (11) and (14), a brief search has been made for a better empirical law. It has been found, as Fig. 3 shows graphically, that a more precise relation is

$$\beta_{c(\text{empir})}=0.415(\beta_{c(\text{Eq. (4)})}+0.455\beta_{c(\text{Eqs. (11), (14)})}), \qquad (15)$$

the notation here being self-explanatory. Equation (15) predicts critical gradients differing from the observed ones by not more than a factor of three, and so is probably the most accurate formula available at the present time.

[4] The authors are indebted to the Ciba Company, Inc., New York, New York, for the gift of oil red C Y dye used in these experiments.

[5] See, e.g., "$\partial T/\partial t$ at the Onset of Convection in Porous Media," by F. T. Rogers, Jr., and H. L. Morrison, read before the 1950 meeting of the Southeastern Section of the American Physical Society in Baton Rouge.

Theory of Axially Slitted Circular and Elliptic Cylinder Antennas

D. R. Rhodes
Department of Electrical Engineering, The Ohio State University Research Foundation, Columbus, Ohio
(Received April 20, 1950)

A method proposed by Sommerfeld for solving boundary value problems involving discontinous surfaces has been applied to the general case of a plane electromagnetic wave of arbitrary direction of incidence and polarization diffracting about one or more sections of perfecting conducting circular and elliptic cylinders of infinite length. The solution is in series form, where the series coefficients are independent for the special case of slots of infinitesimal width (slits). Radiation from the slitted cylinder is restricted to discrete right circular cones about the cylinder axis, each cone corresponding to an ordinary wave-guide mode in the cylinder, while slots of finite width radiate over a continuous range of conical angles. For slots of small but finite width, the relative pattern in the cone $\alpha=\alpha_0$ about a circular cylinder of radius a is the same as in a plane normal to the axis of a cylinder of radius $a\sin\alpha_0$; patterns of the slotted elliptic cylinder are similarly related, where the analog of radius is distance between foci. Conical patterns are shown for the principal TE and TM wave-guide modes in circular cylinders containing one and two diametrically opposed axial slits.

A NUMBER of papers have appeared in the past ten years in which controllable radiation from long axial slots cut in the walls of rectangular wave guide has been reported. In connection with an experimental investigation of radiation from slotted cylinder traveling wave antennas conducted at The Ohio State University[1] it was apparent that a theoretical analysis of this type of antenna would be desirable. A rigorous boundary-value solution involving no assumptions that limit the validity of the solution for a long slot in a rectangular wave guide is, in general, a difficult problem. A slotted cylindrical wave guide whose surface is defined by a single constant coordinate, however, can be treated as a diffraction problem using a method proposed by Sommerfeld[2] in connection with scattering by a cylindrical mirror. Although Sommerfeld applied the method only to two-dimensional scalar fields incident normally on a section of a circular cylinder, the method is easily extended to include the general vector field incident at any angle on elliptic as well as circular cylinders of one or more sections. The solution involves rather lengthy numerical computations in general, but a very simple solution results when the slots are assumed to be narrow. It is the special case of slots of infinite length and vanishing width, or "slits" as they shall be referred to,[3] in a perfectly conducting circular and elliptic cylindrical shell of zero wall thickness which is treated in this paper.

1. DIFFRACTION AROUND SECTIONS OF CIRCULAR CYLINDRICAL SHELLS

The coordinate system of the slotted cylinder is shown in Fig. 1. The fields diffracted about the cylinder by a plane wave incident at an arbitrary angle with respect to the cylinder axis can be obtained by expressing the electric and magnetic fields in terms of two independent single component Hertzian vectors directed along the z axis.[4] Each of the two scalar functions are then solutions of the ordinary scalar wave equation. The required solution is that linear combination of characteristic functions whose coefficients are such that the boundary conditions at the cylindrical surface and at infinity are satisfied.

A single section and two equal diametrically opposed sections of a circular cylinder are shown in cross section

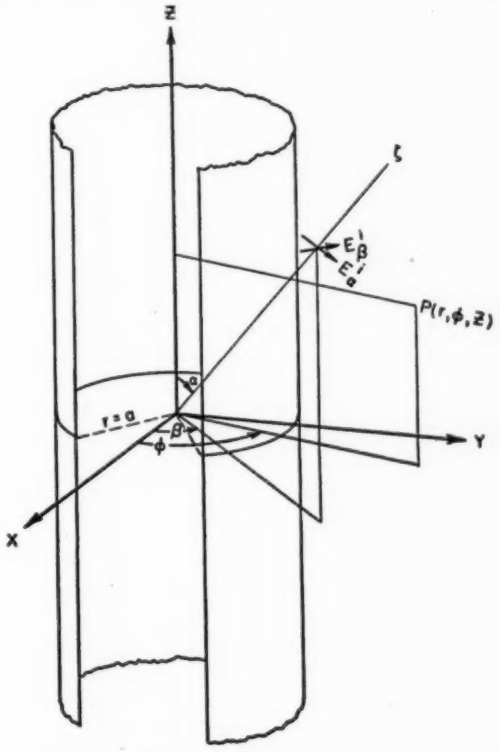


FIG. 1. Coordinate system defining the spherical angles of arrival α and β of the incident plane wave along the ζ-axis and the cylindrical coordinates r, φ, z of an arbitrary point in space.

[1] Hines, Krausz, Rhodes, Rumsey, and Walter, Proc. I.R.E. (to be published).
[2] A. Sommerfeld, *Partial Differential Equations* (Academic Press Inc., New York, 1949), pp. 29–31 and pp. 159–162.
[3] The term "slit" carries the connotation of narrowness to a greater degree than the term "slot" and hence shall be used to specify a slot in this more restrictive sense.
[4] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), pp. 349–51.

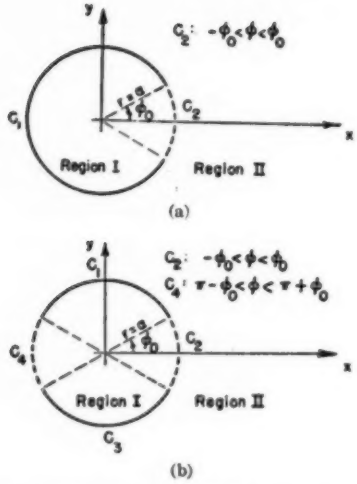


FIG. 2. Cross section of an infinitely long circular cylinder of radius a with (a) a single conducting section, and (b) two equal diametrically opposed conducting sections.

in Fig. 2. These two cases only will be considered. Boundary conditions at the cylindrical surface require that the tangential components of electric field and the normal components of magnetic field vanish over the conducting sections, and that all components of electric and magnetic field match over the aperture separating the interior and exterior regions of the cylinder. The requirement that the scattered radiation consist of outward traveling waves at infinity is satisfied by choosing Hankel functions of the second kind as the radial function in the scattered waves.

For any arbitrary polarization of the incident plane wave it is necessary to use both Hertz vectors. However, if the incident wave is resolved into two orthogonal components, one in and one normal to the plane of incidence, the diffracted field produced by each component can be independently determined using only one Hertz vector. The total field is obtained by superposing the two at every point in space. Elliptically as well as linearly polarized incident waves can be used by including the appropriate time phase between the two orthogonal components.

The incident electric field will, in general, have components $E_\alpha{}^i$ and $E_\beta{}^i$. Each of these components is of the form[5]

$$E^i=E_0e^{ikz\cos\alpha+i\omega t}\sum_{-\infty}^{\infty}i^nJ_n(kr\sin\alpha)e^{in(\varphi-\beta)}. \quad (1)$$

Considering first the component $E_\alpha{}^i$, the fields are derivable from a transverse magnetic Hertzian potential Π_z proportional to E_z. Resolving $E_\alpha{}^i$ into its cylindrical components, the incident potential $\Pi_z{}^i$ is found from $E_z{}^i$ to be

$$\Pi_z{}^i=K\sum_{n=-\infty}^{\infty}i^nJ_n(kr\sin\alpha)e^{in(\varphi-\beta)}, \quad (2)$$

where

$$K=-\frac{E_{0\alpha}}{k^2\sin\alpha}e^{ikz\cos\alpha+i\omega t}. \quad (3)$$

Designating the potential inside the cylinder by $\Pi_z{}^I$ and the potential outside the cylinder by $\Pi_z{}^{II}$, the latter of which can be written as the sum of the incident and scattered potentials

$$\Pi_z{}^{II}=\Pi_z{}^I+\Pi_z{}^s, \quad (4)$$

the potentials can be written in the following form:

$$\Pi_z{}^I=K\sum_{n=-N}^{N}a_nJ_n(kr\sin\alpha)e^{in(\varphi-\beta)}, \quad (5)$$

$$\Pi_z{}^{II}=K\sum_{n=-N}^{N}[i^nJ_n(kr\sin\alpha)+c_nH_n{}^{(2)}(kr\sin\alpha)]e^{in(\varphi-\beta)}, \quad (6)$$

where the summations are limited to a finite number of terms. The exact solution can be approximated as closely as desired by series containing $2N+1$ terms if the arbitrary positive integer N is chosen sufficiently large.

The boundary conditions at $r=a$ must determine the unknown coefficients a_n and c_n uniquely. Three necessary and sufficient independent boundary conditions on the potentials are the following:

$$(\Pi_z{}^I=\Pi_z{}^{II})_{C_1}+_{C_2}, \quad (7a)$$

$$(\Pi_z{}^I=\Pi_z{}^{II}=0)_{C_1}, \quad (7b)$$

$$\left(\frac{\partial\Pi_z{}^I}{\partial r}=\frac{\partial\Pi_z{}^{II}}{\partial r}\right)_{C_2}. \quad (7c)$$

Since the potentials in the two regions must match over the periodic orthogonal interval $0\leq\varphi\leq2\pi$, by Eq. (7a), the coefficients may be equated termwise:[6]

$$a_nJ_n=i^nJ_n+c_nH_n{}^{(2)}. \quad (8)$$

Boundary condition Eq. (7b) requires that

$$\left(\sum_{n=-N}^{N}a_nJ_ne^{in(\varphi-\beta)}=0\right)_{C_1}, \quad (9)$$

and boundary condition Eq. (7c) together with Eq. (8) relating a_n and c_n requires that

$$\left(\sum_{n=-N}^{N}\frac{a_n-i^n}{H_n{}^{(2)}}e^{in(\varphi-\beta)}=0\right)_{C_2}. \quad (10)$$

Equations (9) and (10) relate the coefficients a_n over the two subintervals C_1 and C_2. The solution for the coefficients is referred to by Sommerfeld[2] as a "non-final determination of coefficients," that is, each coefficient depends upon every other coefficient in the series because of the non-orthogonal subintervals. This is a very

[5] Reference 4, p. 371–2. The notation used is essentially that of Stratton with the exception that the time convention has been chosen to be $e^{i\omega t}$ rather than $e^{-i\omega t}$.

[6] The convention will be adopted that all solutions of Bessel's differential equation evaluated at $r=a$ will be written without the argument.

essential distinction from the problem of diffraction around a completely closed conducting surface in which each coefficient is uniquely determined independently of the others. The method of solution proposed by Sommerfeld is that of minimizing the total mean square error incurred over the two subintervals, thereby obtaining the best approximation possible for the finite number of terms used in the series. Since it is required that each summation in Eqs. (9) and (10) is to approximate the null function, the total mean square error over C_1 and C_2 is given by

$$M=\int_{C_1}\left[\sum_{n=-N}^{N} a_n J_n e^{in(\varphi-\beta)}\right]^2 d\varphi + \int_{C_2}\left[\sum_{n=-N}^{N}\frac{a_n-i^n}{H_n^{(2)}} e^{in(\varphi-\beta)}\right]^2 d\varphi. \quad (11)$$

Differentiating M with respect to each unknown coefficient results in the following set of $2N+1$ equations, which must be satisfied simultaneously for the $2N+1$ unknown a_n's:

$$2\pi a_p J_p{}^2 e^{-ip\beta} - \sum_{n=-N}^{N}\left[J_n J_p - \frac{1}{H_n^{(2)}H_p^{(2)}}\right]\mathfrak{a}_{np} e^{-in\beta} a_n = \sum_{n=-N}^{N}\frac{i^n e^{-in\beta}}{H_n^{(2)}H_p^{(2)}}\mathfrak{a}_{np}, \quad |p|\leq N, \quad (12)$$

where δ_{np} is Kronecker's delta and where the integral of products of angular functions over C_2 is given by

$$\mathfrak{a}_{np}=\frac{2\sin(n-p)\varphi_0}{n-p}, \quad (13)$$

$\mathfrak{a}_{pp}$ being defined as the limit when $n\to p$.

The analysis for transverse electric fields excited by the other component $E_\beta{}^i$ of the incident plane wave is exactly analogous to the preceding analysis for transverse magnetic fields, where K is replaced by

$$K^*=-\frac{E_{0\beta}}{\eta k^2\sin\alpha} e^{ikz\cos\alpha+i\omega t},$$

where $\eta=(\mu/\epsilon)^{\frac{1}{2}}$. The equations relating coefficients $a_n{}^*$ and $c_n{}^*$ correspond to those relating a_n and c_n in Eqs. (8) and (12) when the radial functions are replaced by their derivatives.

The fields diffracted around a multiple section cylindrical shell by an incident plane wave are treated in the same manner as the single section cylinder. As with the single-section cylinder the boundary conditions must be satisfied over each of the conducting sections and apertures between sections. Minimizing the total mean square error over the cylinder again gives the best approximation in the sense of least squares. Although the method is applied here only to the case of two equal diametrically opposed sections, it can be easily extended to include more than two sections.

For the transverse magnetic component $E_\alpha{}^i$ of a plane wave incident on the cylinder shown in cross section in Fig. 2b, boundary conditions again require that a_n and c_n be related by Eq. (8) and that Eqs. (9) and (10) must hold over the subintervals C_1, C_3 and C_2, C_4, respectively. The resulting relationship between coefficients a_n then reduces to the following set of equations:

$$\pi a_p J_p{}^2 e^{-ip\beta} - \sum_{n=-N}^{N}{}'\left[J_n J_p - \frac{1}{H_n^{(2)}H_p^{(2)}}\right]\mathfrak{a}_{np} e^{-ip\beta} a_n = \sum_{n=-N}^{N}{}'\frac{i^n e^{-in\beta}}{H_n^{(2)}H_p^{(2)}}\mathfrak{a}_{np}, \quad |p|\leq N, \quad (14)$$

where the primed summation is on even values of n when p is even and odd values of n when p is odd. The set of equations relating coefficients $a_n{}^*$ for the transverse electric component $E_\beta{}^i$ correspond to Eqs. (14) when the radial functions are replaced by their derivatives.

It can be seen from Eq. (12) that the coefficients a_n are functions of the angles α and β of the incoming plane wave. This is a severe limitation when solving for the pattern of the scattered field or of a detecting probe near the conducting surface, since it is necessary to repeat the solution of the system of equations for each angle of incidence. The limitation vanishes, however, for the special case of a vanishingly narrow slot in a cylinder.

2. SERIES COEFFICIENTS FOR NARROW SLOTS

As the slot width vanishes, the set of Eqs. (12) relating coefficients a_n reduces to a set of independent equations, each equation involving only a single coefficient, since C_1 approaches the interval of orthogonality of the angular functions. When the angular slot width $2\varphi_0$ is small, the integrals over C_1 are essentially independent of n and p, reducing to

$$\mathfrak{a}_{np}\approx 2\varphi_0. \quad (15)$$

Consider first the transverse magnetic component of the plane wave incident on the cylinder. When the argument $ka\sin\alpha$ is not a zero of J_q, then by choosing φ_0 vanishingly small all of the terms in the summation on a_n become vanishingly small with respect to the unsummed term containing a_p, Eqs. (12) then becoming

$$2\pi a_p J_p{}^2 e^{-ip\beta}=\frac{2\varphi_0}{H_p^{(2)}}F_c(\alpha,\beta), \quad |p|\leq N, \quad (16)$$

where

$$F_c(\alpha,\beta)=\sum_{n=-N}^{N}\frac{i^n e^{-in\beta}}{H_n^{(2)}}. \quad (17)$$

When $ka\sin\alpha$ is a zero of J_q, then for $p\neq q$ the set of Eqs. (12) reduce as before to Eq. (16). For $p=q$, however, the unsummed term containing a_q is zero and all

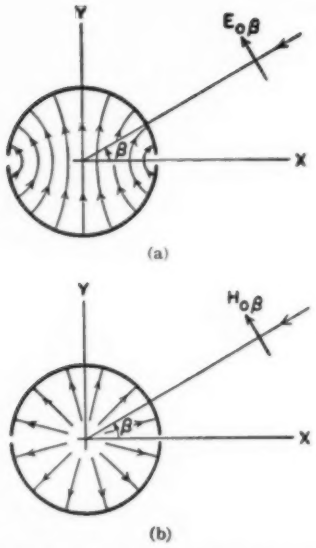


FIG. 3. Electric field distribution in the xy-plane within a slitted cylinder excited by a plane wave incident at (a) α_{11}', and (b) α_{01}.

of the a_n's in the summation are vanishingly small except the one containing a_q; hence at $p=q$ the set of Eqs. (12) reduces to

$$a_q e^{-iq\beta} = -iN_q F_c(\alpha, \beta), \quad |q| \leq N, \tag{18}$$

which determines each coefficient a_q explicitly. The coefficients a_n* for transverse electric excitation correspond as before when the radial functions are replaced by their derivatives. A similar solution follows immediately from Eq. (14) for the double section cylinder.

The scattered radiation reduces to that of a completely closed cylinder in the limit as φ_0 vanishes, since a_n vanishes everywhere except at the zeros of J_q. For a_n to vanish everywhere including the zeros of J_q it is only necessary to let φ_0 vanish before J_q passes through its zeros. The coefficients for the scattered radiation are then obtained directly from Eq. (8).

3. MODES IN THE SLITTED CIRCULAR CYLINDER

a. Existence of Wave-Guide Modes at Certain Critical Incident Angles

The electric and magnetic fields everywhere inside a cylinder containing one or more slits have vanishingly small amplitudes due to the factor φ_0, except when the incident plane wave arrives at such an angle that

$$J_q(ka \sin\alpha_{qm}) = 0, \tag{19a}$$

or

$$J_q'(ka \sin\alpha_{qm}') = 0, \tag{19b}$$

where α_{qm} and α_{qm}' denote the incident angles for which the mth zeros of J_q and J_q' occur, respectively. The series representing each field component then reduces to two terms, namely, those for which $n=\pm q$. At α_{qm} the components of electric field in a cylinder containing a single slit reduce to the following:

$$E_r^{(1)} = -\epsilon_q E_{0\alpha}[\cos\alpha_{qm} J_q'(kr \sin\alpha_{qm}) \times \cos q\varphi] N_q F_c(\alpha, \beta), \tag{20a}$$

$$E_\varphi^{(1)} = \epsilon_q E_{0\alpha}\left[\cos\alpha_{qm} \frac{q}{kr \sin\alpha_{qm}} \times J_q(kr \sin\alpha_{qm}) \sin q\varphi\right] N_q F_c(\alpha, \beta). \tag{20b}$$

$$E_z^{(1)} = i\epsilon_q E_{0\alpha}[\sin\alpha_{qm} J_q(kr \sin\alpha_{qm}) \times \sin q\varphi] N_q F_c(\alpha, \beta), \tag{20c}$$

where ϵ_q is Newman's number; the common factor $\exp(ikz\cos\alpha_{qm} + i\omega t)$ has been omitted. Equations (20) are precisely the expressions given by Stratton[7] for the electric field components of the TM_{qm} mode in a hollow circular cylinder wave guide if Stratton's coefficients a_{qm} are chosen to be

$$a_{qm} = i\epsilon_q E_{0\alpha}(\lambda/\lambda_{qm}) N_q F_c(\alpha, \beta),$$

and if the critical angle α_{qm} is related to the cut-off wave-length λ_{qm} of the TM_{qm} circular wave-guide mode in such a way that

$$\sin\alpha_{qm} = \lambda/\lambda_{qm} = u_{qm}/ka, \tag{21}$$

where u_{qm} is the mth zero of J_q. Similarly, it can be shown that a plane wave incident at $\alpha=\alpha_{qm}'$ excites the TE_{qm} mode in a slitted circular cylinder, where α_{qm}' is related to the wave-length of the exciting plane wave and to its cut-off wave-length by Eq. (21). The field distribution of the lowest order TE and TM modes excited in a slitted cylinder are shown in Fig. 3.

An interesting philosophical but physically useless observation is the fact that a non-vanishing energy bearing field can exist in the cylinder in the ideal case of a plane wave incident at the critical conical angles on a perfectly conducting cylindrical shell, even when the shell is completely closed. These fields are physically unobtainable, since they must collapse if any energy is withdrawn from the system.

b. Modes Dependent on Slit Width

At all angles of incidence other than those in the cones α_{qm} and α_{qm}' the coefficients of the series for the fields in the cylinder are proportional to the slit width. Even at the critical angles defined by these cones all but one pair of coefficients are proportional to slit width, the fields associated with this pair being identified with ordinary wave-guide modes in the cylinder. As a consequence, for incidence at angles other than the critical angles all fields in the cylinder vanish with slit width. Since there has been some interest expressed by various

[7] Reference 4, p. 541, Eqs. (75). A typographical error is apparent in the angular function in the expression for E_z.

authors[8–10] in one special case in which these vanishing modes are excited, namely, the case of a plane wave incident normally on a slitted cylinder of arbitrary radius, these modes will be briefly considered.

For a single slit in a cylinder, the axial component of electric field for TM waves and the angular component of electric field for TE waves, for incidence at angles other than α_{qm} and α_{qm}', are given by the following:

$$E_z^{(1)} = -\frac{\varphi_0}{\pi} E_{0\alpha} \sum_{p=0}^{N} \frac{\epsilon_p}{J_p^2 H_p^{(2)}} \times [\sin\alpha J_p(kr\sin\alpha)\cos p\varphi] F_c(\alpha,\beta), \quad (22a)$$

$$E_\varphi^{(2)} = -i\frac{\varphi_0}{\pi} E_{0\beta} \sum_{p=0}^{N} \frac{\epsilon_p}{J_p'^2 H_p^{(2)\prime}} \times [J_p'(kr\sin\alpha)\cos p\varphi] F_c'(\alpha,\beta), \quad (22b)$$

where $F_c'(\alpha,\beta)$ corresponds to $F_c(\alpha,\beta)$ when the radial functions are replaced by their derivatives. For a cylinder containing two diametrically opposed slits the conical pattern factors $F_c(\alpha,\beta)$ and $F_c'(\alpha,\beta)$ are each summed over even values of n when p is even and odd values of n when p is odd. Individual terms in the summation on p in Eqs. (22) do not represent ordinary wave-guide modes, since no relation such as that given by Eq. (21) can exist for arbitrary values of α.

At normal incidence, Eqs. (22) contain the same conical pattern factors which Sinclair[8] obtained in a plane normal to the cylinder axis for a very narrow axial slot in a cylinder when considered as a transmitting antenna. Sinclair resolved an assumed uniformly distributed tangential electric field at $r=a$ into a Fourier series and equated this series termwise to a second series representing the total electric field for $r\geq a$. Using Sinclair's method, Papas and King[9] computed the current distribution on the surface of a circular cylinder excited by an electric field across an axial slit. Sinclair has also shown[10] that the pattern, in a plane normal to the cylinder axis, of arrays of axial slits on the surface of conducting circular cylinders can be calculated by superposing the fields caused by each slit. The patterns of two diametrically opposed slits obtained by using Sinclair's method of arraying slits are the same as given here.

4. OPERATION AS A TRAVELING WAVE ANTENNA

Operation of the slitted cylinder as a receiving antenna is readily apparent. As the conical angle α is continuously varied, the fields in the cylinder abruptly change at α_{qm} and α_{qm}' from fields having vanishingly small amplitudes (proportional to the slit width) into a single TM or TE wave-guide mode whose amplitude is independent of the slit width. Since the zeros of J_q and J_q' are distinct, no two wave-guide modes are excited at the same angle of incidence and hence each mode is distinct (exceptions are the TE_{0m} and TM_{1m} modes, which are excited at the same incident angle since $J_0' = -J_1$). A receiving probe placed in the cylinder has an open-circuit voltage induced in it proportional to the field at the probe. A slotted cylinder antenna should exhibit directional characteristics dependent on slot width, approaching infinite directivity as the slot width vanishes.

The critical conical angles were determined several years ago by Hansen[11] by considering axially slitted wave guides as transmitting traveling wave antennas. Maximum radiation occurs at those angles for which radiation from every infinitesimal element of the slit aperture interferes constructively. Referring to Fig. 4, maximum constructive interference occurs when α_{qm} or α_{qm}' is related to the excitation wave-length and to the guide wave-length by

$$\cos\alpha_{qm} = \lambda/\lambda_g. \quad (23)$$

Equation (23), obtained from simple interference considerations, is equivalent to Eq. (21) obtained by rigorous solution of the boundary value problem.

The critical cones associated with each wave-guide mode are distinguished by the roots u_{qm} and u_{qm}'. For any given wave-guide mode these cones are related to the electrical radius of the cylinder by Eq. (21). The angles α_{qm} and α_{qm}' are plotted in Fig. 5 as a function of the electrical radius a/λ for cylinders not exceeding a radius of one wave-length. The critical cones asymptoti-

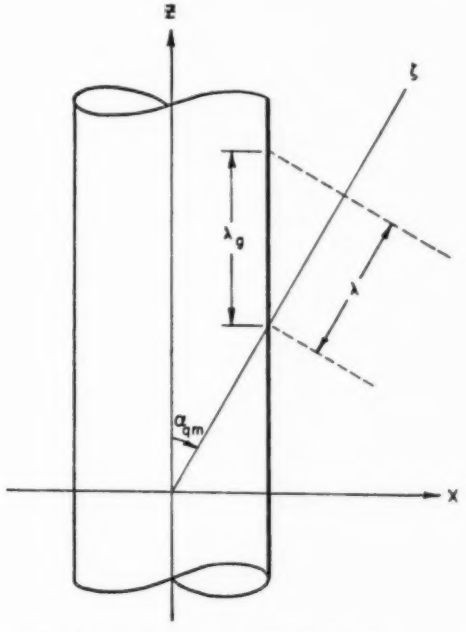


FIG. 4. The slitted cylinder as a traveling wave antenna. Maximum constructive interference occurs in cones defined by $\cos\alpha_{qm} = \lambda/\lambda_g$, where λ is free space wave-length and λ_g is guide wave-length.

[8] Sinclair, Jordan, and Vaughn, Proc. I.R.E. 35, 1451–62 (1947).
[9] C. H. Papas and R. King, Quart. App. Math. 7, 175–82 (1949).
[10] George Sinclair, Proc. I.R.E. 36, 1487–92 (1948).
[11] W. W. Hansen, U. S. Patent No. 2402622.

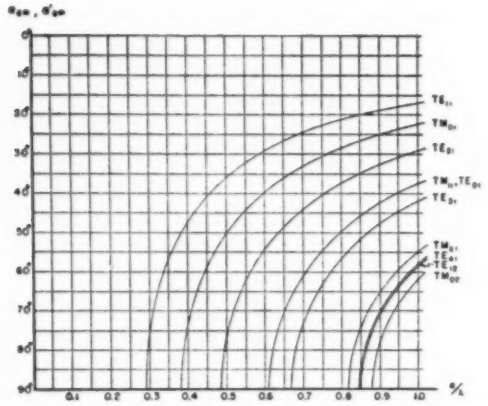


FIG. 5. Critical conical angles α_{qm} and α_{qm}' as a function of electrical radius of the slitted circular cylinder.

cally approach the cylinder axis as the excitation frequency is increased. In contrast to the unbounded upper frequency limit, each mode has a sharply defined low frequency limit. The low frequency limit occurs when maximum radiation is in the plane normal to the cylinder axis, and corresponds to the condition of cut-off in the cylindrical wave guide. For all frequencies below the low frequency limit the wave guide is operating below its cut-off frequency, and the energy propagating inside the cylinder is rapidly attenuated.

The slitted cylinder antenna is fundamentally linearly polarized, the electric field radiated by TM modes being E_α and that by TE modes being E_β.

The theory of the slitted cylinder can be used to predict the conical pattern and critical cones of wave-guide-fed slits in cylinders. The critical cones are determined by the phase velocity of waves traveling in the slit, which is determined entirely by the wave-guide cross-section; the conical pattern depends on the cross sections of both wave guide and cylinder. For the case of a circular wave guide feeding a slotted circular cylinder as shown in Fig. 6a, the argument of the radial functions in the conical pattern factor $F_c(\alpha_{qm}, \beta)$ for the TM_{qm} mode in the wave-guide is

$$kb \sin\alpha_{qm} = (b/a)u_{qm}. \tag{24}$$

Two or more wave-guide-fed slits in a cylinder, as suggested in Fig. 6b, may be independently excited and the radiation pattern calculated by Sinclair's method of superposing fields.

5. RADIATION PATTERNS

The relative conical radiation patterns of the slitted cylinder are plots of the conical pattern factors $F_c(\alpha, \beta)$ and $F_c'(\alpha, \beta)$. The pattern factors are complex functions of β from which the amplitude and phase patterns can be obtained. The conical pattern is fixed for any given wave-guide mode in a slitted cylinder regardless of the cylinder radius or critical conical angle. In Fig. 7 is shown the conical amplitude patterns characteristic of the lowest order TE and TM modes in a circular cylinder containing one or two diametrically opposed slits. An example of the three dimensional pattern of a slitted cylinder is shown in Fig. 8 for a $\frac{7}{8}\lambda$ diameter cylinder containing one or two diametrically opposed slits. The patterns in the critical cones are those shown in Fig. 7 when plotted on an absolute scale. The critical cones for the cylinder diameter chosen are $\alpha_{11}' = 42°$ and $\alpha_{01} = 61°$.

Since the incident angle α appears in the pattern factor only in the product $ka \sin\alpha$, it follows that the relative pattern in any cone $\alpha = \alpha_0$ about a cylinder of radius $r = a$ is the same as the relative pattern in a plane normal to the axis of a cylinder of radius $r = a \sin\alpha_0$.

6. SLITTED ELLIPTIC CYLINDER

The preceding method outlined for the slitted circular cylinder can be applied equally well to the slitted elliptic cylinder, the only other closed cylindrical surface defined by a single coordinate for which the scalar wave equation is separable. Elliptic cylinder coordinates (u, v, z) are related to cartesian coordinates (x, y, z) by the transformation

$$x = c_0 \cosh u \cos v, \quad y = c_0 \sinh u \sin v;$$

the metric coefficients are given by $h_u = h_v = c_0(\cosh^2 u - \cos^2 v)^{\frac{1}{2}}$. The coordinate system in Fig. 9 shows a conducting elliptic cylinder defined by $u = u_0$ with a slot of angular width $2v_1$, located at $v = v_0$. The coordinates u and v are radial and angular coordinates corresponding to r and φ of the circular cylinder.

The wave functions resulting from separation of the wave equation consist of products of radial and angular Mathieu functions. Using the series expansion of a plane wave in terms of even and odd elliptic wave functions as given by Stratton,[12] the transverse magnetic potentials inside and outside the cylinder $u = u_0$ can be written as follows:

$$\begin{aligned}\Pi_z^I = (8\pi)^{\frac{1}{2}} K \sum_{n=0}^{N} [&(a_n^e/N_n^e) Je_n(c, \cosh u) \\ &\times Se_n(c, \cos v) Se_n(c, \cos\beta) \\ &+ (a_n^o/N_n^o) Jo_n(c, \cosh u) \\ &\times So_n(c, \cos v) So_n(c, \cos\beta)]. \end{aligned} \tag{25a}$$

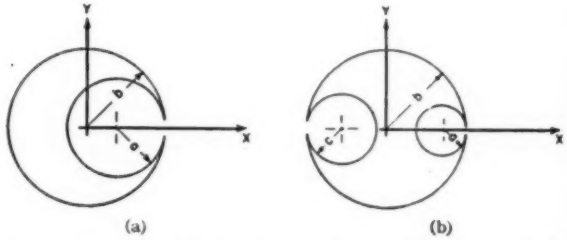


FIG. 6. Wave-guide-fed slits in cylinders. Patterns are calculable from slitted cylinder theory.

[12] Reference 4, p. 386, Eq. (84).

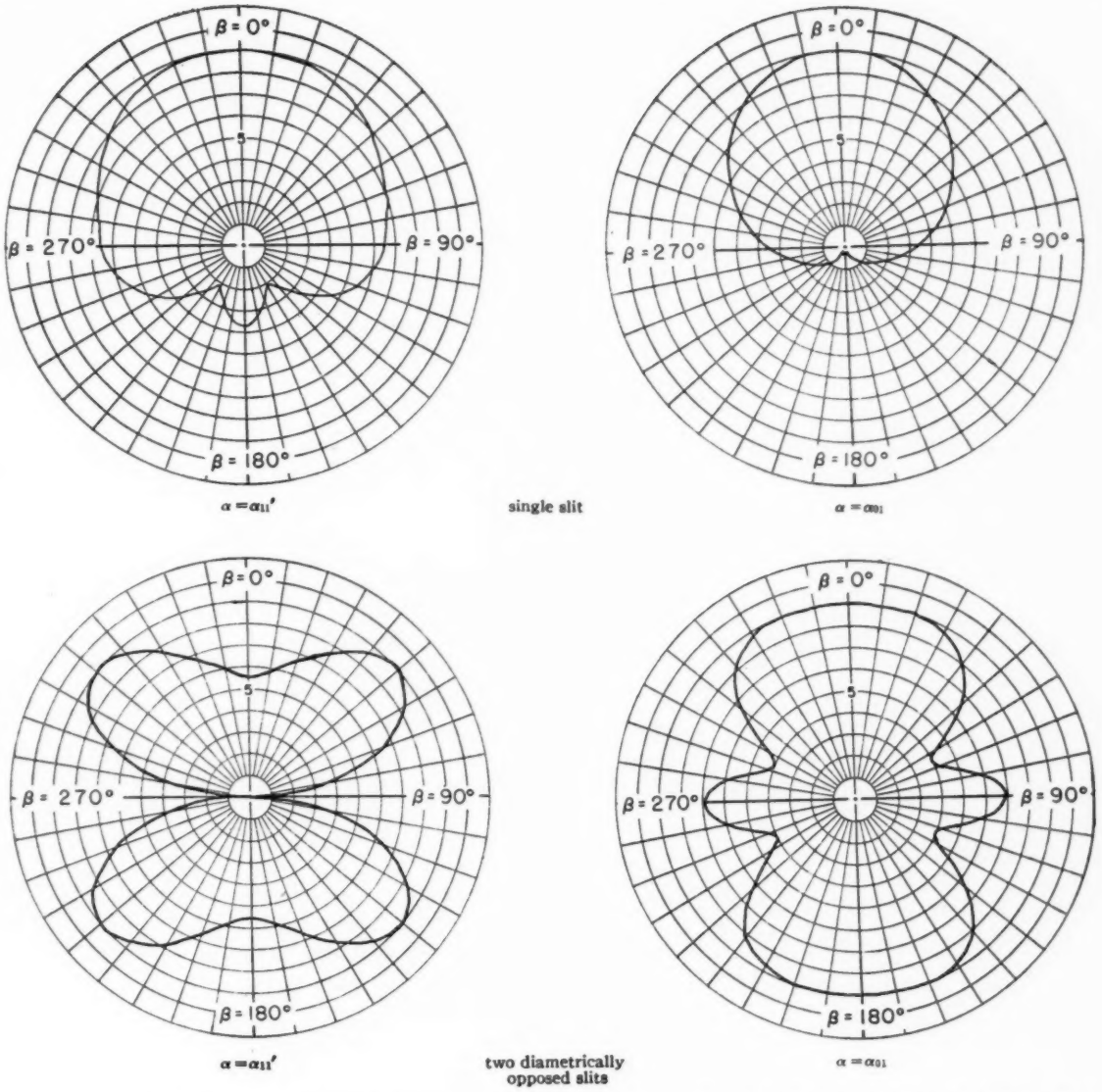


FIG. 7. Relative conical patterns of a circular cylinder containing one or two diametrically opposed slits when the cylinder is excited in the TE_{11} or TM_{01} circular wave-guide mode.

$$\Pi_z^{II}=(8\pi)^{\frac{1}{2}}K\sum_{n=0}^{N}\left[\frac{1}{N_n^e}\{i^nJe_n(c,\cosh u)+c_n^eHe_n^{(2)}(c,\cosh u)\}\times Se_n(c,\cos v)Se_n(c,\cos\beta)+(1/N_n^o)\{i^nJo_n(c,\cosh u)+c_n^oHo_n^{(2)}(c,\cosh u)\}\times So_n(c,\cos v)So_n(c,\cos\beta)\right]. \quad (25b)$$

The boundary conditions require that Π_z match over the interval $0\leq v\leq 2\pi$ and vanish over C_1, and that $\partial\Pi_z/\partial u$ match over C_2. The condition that Π_z must match over the interval of orthogonality of the angular functions requires that

$$a_n^eJe_n=i^nJe_n+c_n^eHe_n^{(2)}, \quad (26a)$$

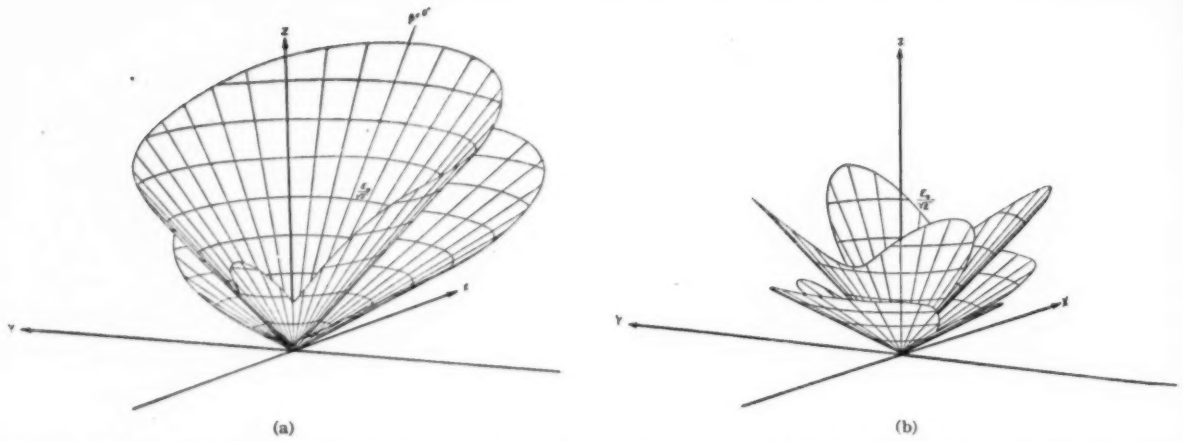
$$a_n^oJo_n=i^nJo_n+c_n^oHo_n^{(2)}, \quad (26b)$$

where the radial functions are evaluated at $u=u_0$. The other two boundary conditions relate the coefficients over the non-orthogonal subintervals C_1 and C_2. Minimizing the total mean square error and letting v_1 vanish again results in a set of independent equations for the unknown coefficients a_n. When

$$Je_{2e}(c,\cosh u_0)=0, \quad (27a)$$

or

$$Jo_{2o}(c,\cosh u_0)=0, \quad (27b)$$



(a) (b)

FIG. 8. Three-dimensional pattern of a radial electric probe located at $\varphi=(\pi/2)$ in a $\frac{3}{4}\lambda$ diameter cylinder containing (a) a single slit located at $\varphi=0$, and (b) two diametrically opposed slits located at $\varphi=0$ and $\varphi=\pi$. The incident wave is polarized such that $E_{0\alpha}=E_{0\beta}=(\sqrt{2}/2)E_0$. The patterns in the critical cones are taken from Fig. 7. Patterns at $\alpha=\alpha_{01}$ are shown multiplied by a factor of two.

the coefficients for $p\neq q$ are given by

$$a_p{}^e Je_p{}^2 Se_p{}^\beta = \frac{2v_1 Se_p{}^{v_0}}{He_p{}^{(2)}} F_e(c, v_0, \beta), \quad p \leq N, \quad (28a)$$

$$a_p{}^o Jo_p{}^2 So_p{}^\beta = \frac{2v_1 So_p{}^{v_0}}{Ho_p{}^{(2)}} F_e(c, v_0, \beta), \quad p \leq N, \quad (28b)$$

while that for $p=q_e$ is given by

$$a_{q_e}{}^e \frac{Se_{q_e}{}^\beta}{Ne_{q_e}} = -i \frac{Ne_{q_e}}{Se_{q_e}{}^{v_0}} F_e(c, v_0, \beta), \quad q_e \leq N, \quad (29a)$$

and that for $p=q_o$ is given by

$$a_{q_o}{}^o \frac{So_{q_o}{}^\beta}{No_{q_o}} = -i \frac{No_{q_o}}{So_{q_o}{}^{v_0}} F_e(c, v_0, \beta), \quad q_o \leq N, \quad (29b)$$

where $Se_{q_e}{}^\beta$ represents $Se_{q_e}(c, \cos\beta)$, $Se_{q_e}{}^{v_0}$ represents $Se_{q_e}(c, \cos v_0)$, etc., and where the elliptic conical pattern factor is given by

$$F_e(c, v_0, \beta) = \sum_{n=0}^{N} i^n \left(\frac{Se_n{}^\beta Se_n{}^{v_0}}{N_n{}^e He_n{}^{(2)}} + \frac{So_n{}^\beta So_n{}^{v_0}}{N_n{}^o Ho_n{}^{(2)}} \right). \quad (30)$$

The coefficients $a_n{}^{e*}$ and $a_n{}^{o*}$ for TE modes are obtained from those for TM modes by replacing the radial functions by their derivatives in Eqs. (28) to (30).

The fields in the cylinder at the critical angles ${}_e\alpha_{qm}$ and ${}_o\alpha_{qm}$ are the ordinary wave-guide modes in the elliptic wave guide.[13] At all angles other than the critical angles the fields vanish according to Eq. (28). The elliptic conical pattern factor in a plane normal to the axis is the same as that derived by Sinclair[14] and Carter[15] for the two-dimensional TE case. Since the incident angle α appears in the pattern factor only in the product $c=kc_0\sin\alpha$, the pattern in any cone $\alpha=\alpha_0$ about a cylinder $u=u_0$ whose distance between foci is $2c_0$ is the same as the pattern in a plane normal to the axis of a cylinder $u=u_0$ whose distance between foci is $2c_0\sin\alpha_0$.

The transition to the circular case follows as the limit when $c_0\to 0$ and $u\to\infty$, for then $c_0\cosh u\to r$. The solution for the elliptic cylinder then goes over to that for the circular cylinder.

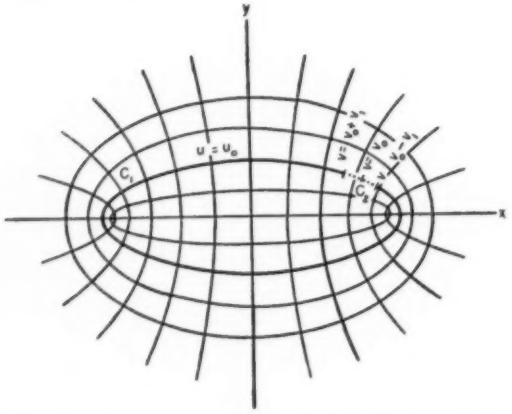


FIG. 9. Coordinate system of the slitted elliptic cylinder. The cylinder is defined by $u=u_0$ with distance between foci $2c_0$. A single slit is located at $v=v_0$ with angular width $2v_1$.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to George Sinclair for his introduction to the theory of diffraction problems, and to V. H. Rumsey for his interest and constructive criticisms.

Work described in this paper was carried out, in part, under a contract between the Wright-Patterson Air Force Base of the Air Materiel Command and The Ohio State University Research Foundation.

[13] L. J. Chu, J. App. Phys. **9**, 583–91 (1938).

[14] George Sinclair, "The Patterns of Antennas Mounted near Elliptic Cylinders," The Ohio State University Research Foundation Report No. 301–1, June, 1947.

[15] Carter, Martin, and Thorne, "Slot Antennas in Elliptic Cylinders and Airplane Vertical Stabilizer," RCA Laboratories Report No. CM-45-15, April, 1948.

Frequency Factor and Activation Energy for the Volume Diffusion of Metals*

G. J. DIENES
Atomic Energy Research Department, North American Aviation, Inc., Los Angeles, California
(Received May 12, 1950)

The theoretical equations proposed for calculating D_0 of the diffusion-constant-temperature relation

$$D=D_0e^{-E/RT}$$

are critically discussed in the light of presently available data. It is concluded that none of the relations gives satisfactory agreement with experiment.

Empirical analysis of experimental data covering a wide range of D_0 and E values showed that the quantity E/T_m, where T_m is the melting point in °K, is the main factor in determining the value of D_0, and that D_0 depends approximately exponentially on E/T_m. In addition to many other diffusion constants, it is shown that the anisotropic diffusion of zinc and the widely differing diffusion constants for self-diffusion in α- and γ-iron are included in the proposed empirical correlation.

A possible theoretical basis for the empirical correlation is described based on the idea of local melting or disordering. The quantity E/T_m is identified with the entropy of activation.

I. INTRODUCTION

IT is well known that the temperature dependence of diffusion can be well described by the simple exponential relation

$$D=D_0e^{-E/RT}, \tag{1}$$

where D=diffusion constant; D_0=constant, same dimensions as D; E=activation energy; R=gas constant; T=temperature in °K. Many attempts have been made to derive theoretical or semi-empirical expression for D_0, the so-called frequency factor. Recently, sufficient data on volume diffusion have become available to judge critically the validity of these equations.[1]

It may be pointed out that investigations over the last ten years or so have shown very clearly that the diffusion constant as well as the activation energy is a function of the concentration in most intermetallic systems. The importance of proper thermodynamic treatment has been emphasized (use of activities in place of concentrations). A review by A. D. Le Claire[2] summarizes well our present state of knowledge, and it is clear that the corresponding D_0 of such systems is not yet amenable to theoretical analysis. Any relation proposed for D_0, therefore, should be tested by comparison with self-diffusion data and with intermetallic diffusion data obtained at very low concentrations and far below the solubility limits so that concentration dependence is negligible. Since grain boundary and surface diffusion present special problems, the present discussion will be limited to volume diffusion.

It should also be remarked that evaluation of D_0 should not be tested by comparing experimental and calculated values of the activation energy, E. Such a calculation is found to be quite insensitive to the value of D_0 chosen because of the masking effect of the exponential factor in Eq. (1). D_0 values now known cover a sufficiently wide range for proper comparison of any proposed relation with experiment.

Table I is a summary of proposed theoretical and semi-empirical relations. (For notation, see headings of Table II.) In the third column the implied linear functional relations are also indicated. Leichter[3] has pointed out that only the Langmuir-Dushman[4] equation and the relation proposed by him predict the right order of magnitude for D_0. This conclusion is based on a few self-diffusion studies.[3] It is probably far more important, however, that none of the relations listed in Table I predict the correct functional dependence. Sufficient data are given in Table II for systems of known self-diffusion to verify this statement by simply plotting the data according to column three of Table I. Among others, glaring examples of deviations may be found in the anisotropy of Zn and in the self-diffusion of iron in the α- and γ-phase.[5]

Bradley's[6] relation comes closest to showing a correlation provided the relation between D_0 and $E\nu d^2/T_M$ (where T_M is the mean temperature of the experiment) is taken to be an exponential rather than a linear one. It will become clear from later discussion in this paper that the above is to be expected for the simple reason that T_M is usually not too far removed from the melting point.

* This document is based on work performed under Contract No. AT-11-1-GEN-8 for the AEC.

[1] C. J. Smithells, *Metals Reference Book* (Interscience Publishers, Inc., New York, 1949), pp. 390–412.

[2] A. D. Le Claire, Section on Diffusion of Metals in *Progress in Metal Physics I* (Interscience Publishers, Inc., New York, 1949), pp. 306–379.

For surveys of the field of diffusion in metals see:

R. M. Barrer, *Diffusion in and Through Solids* (Cambridge University Press, 1941).

W. Jost, *Diffusion und Chemische Reaktion in Festen Stoffen* (Theodor Steinkopff Verlag., Dresden, 1937).

W. Seith, *Diffusion in Metallen* (Verlag. Julius Springer, Berlin, 1939).

F. Seitz, *The Physics of Metals* (McGraw-Hill Book Company, Inc., New York, 1943).

[3] M. Leichter, N.A.C.A. Tech. Note No. 1856, Washington, April, 1949.

[4] I. Langmuir and S. Dushman, Phys. Rev. **20**, 113 (1922).

[5] C. F. Birchenall and R. F. Mehl, J. App. Phys. **19**, 217 (1948).

[6] R. S. Bradley, Trans. Faraday Soc. **33**, 1185 (1937).

TABLE I. Equations proposed for D_0 of $D=D_0e^{-E/RT}$.

Investigator	Proposed equation	Implied linear functional relation
Langmuir-Dushman[a]	$D_0=(E/Nh)d^2$	D_0 *vs.* Ed^2
Bradley[b]	$D_0=1/6(E/RT)\nu d^2$	D_0 *vs.* $E\nu d^2/T_m$
Van Liempt[c]	$D_0=8\nu d^2/3\pi$	D_0 *vs.* νd^2
Eyring[d]	$D_0=(V/N)^{\frac{1}{3}}(KT/h)(1-e^{-3\theta/4T})$	D_0 *vs.* $V^{\frac{1}{3}}\theta$*
Leichter[e]	$D_0/\nu d^2=2\pi W/3(3E_1/E_t)^{\frac{1}{2}}e^{E_f/2RT}$	$\mathrm{Log}D_0/\nu d^2$ *vs.* E_f/T_m**
Cichocki[f]	$D_0=2.43\times10^4(M)^{\frac{1}{2}}/V(T_m)^{\frac{1}{2}}$	D_0 *vs.* $(M)^{\frac{1}{2}}/V(T_m)^{\frac{1}{2}}$

* Exponential can be expanded for all values of interest here.
** Other quantities are essentially constant for values of interest here.
[a] See reference 4.
[b] See reference 6.
[c] J. Van Liempt, Zeits. f. Physik **96**, 534 (1935).
[d] See reference 11.
[e] See reference 3.
[f] J. Cichocki, J. de phys. et rad. **7**, 420 (1936).

II. EMPIRICAL CORRELATION

The discussion given so far indicates that no suitable expression exists (theoretical or empirical) for relating D_0 to, and for calculating it from, other physical quantities. An analysis of presently available experimental data was undertaken in order to find out whether it is possible to abstract from the data themselves the physically significant quantities which determine D_0. Guided by the often proposed idea of melting, or disorder, as an important factor in diffusion, it was empirically discovered that the physically important quantity is

$$E/T_m,$$

where E=activation energy; T_m=melting point in °K. This conclusion is based on the data given in Tables II and III. An attempt was made to include all reliable diffusion data within the restrictions discussed in the introduction. In cases where several different values are available, Smithell's judgment has been accepted in most cases as to reliability (as given in his tables).[1] It was found that good correlation is obtained over a wide range of D_0 values by plotting $\log(D_0/\nu d^2)$ or $\log(D_0/\nu\lambda^2)$ against E/T_m. The best correlation appears to be given by $\log(D_0/\nu\lambda^2)$ *vs.* E/T_m. The possibility of a close relation between E and T_m for self-diffusion has been discussed by many workers.[7] It is to be emphasized that the present discussion is concerned with the relation between D_0 and the above quantities, rather than the possible interrelation of E and T_m.

The results are shown in Fig. 1, where $\log(D_0/\nu\lambda^2)$ is plotted against $E/2.303RT_m$ for all the items of Table II and III. Several values are available for Cu with three typical ones given in Table II. All of these fall along the correlation curve and furnish no judgment as to the best values for this material. This is probably due to compensating errors in E and D_0 in the various experimental data. In Table III all physical constants refer to the solvent atoms. This is obviously incorrect in the immediate neighborhood of a solute atom but no data are available for more accurate analysis. These variations must be minor, however, relative to the exponential influence of E/T_m. Within the experimental reliability of D_0 values the correlation is seen to be a very good one and covers a D_0 range from about 10^{-9} to 10^{+4} and activation energies from 8 to 140 kcal./mole. This curve clearly establishes E/T_m as the main factor which determines the value of D_0. However, since the influence of this factor is approximately an exponential

TABLE II. Data for self-diffusion of metals.[a]

System	Frequency factor, D_0, $\mathrm{cm^2\ sec.^{-1}}$	$\log\frac{D_0}{\nu d^2}$	$\log\frac{D_0}{\nu\lambda^2}$	Activation energy, E, kcal./mole	Melting point, T_m, K°	$\frac{E}{2.303RT_m}$	Lattice constant, d, cm $\times10^{-8}$	Interatomic distance, λ, cm $\times10^{-8}$	Debye frequency, ν, $\mathrm{sec.^{-1}}$ $\times10^{12}$	Debye temp. θ	Av. T. of exp. T_m °K	Molar volume, V, $\mathrm{cm^3}$	Heat of fusion E_f kcal./mole
Silver	8.95×10^{-1}	2.08	2.38	45.9	1234	8.32	4.08	2.88	4.5	215	1110	10.8	2.73
Gold	1.6×10^{-1}	1.41	1.71	53.0	1336	8.68	4.07	2.87	3.8	170	1183	10.4	3.06
Copper	4.7×10^{1}	3.74	4.04	61.4	1357	9.89	3.61	2.55	6.6	315	1113	7.1	3.11
Copper[b]	1.1×10^{1}	3.11	3.41	57.0	1357	9.20	3.61	2.55	6.6	315	1113	7.1	3.11
Copper[c]	3.0×10^{-1}	1.54	1.84	46.8	1357	7.55	3.61	2.55	6.6	315	1113	7.1	3.11
Lead	6.6	3.17	3.48	27.9	600	10.15	4.94	3.48	1.8	88	488	18.3	1.19
Tungsten	1.15×10^{1}	3.24	3.37	140.0	3655	8.37	3.16	2.73	6.6	310	~3000	9.55	11.2
Zinc, (||) c-axis	4.6×10^{-2}	0.55	0.55	20.4	693	6.44	4.94	4.94	5.3	250	655	9.2	1.72
Zinc, ($\perp$) c-axis	9.1×10^{1}	4.39	4.39	31.0	693	9.77	2.65	2.65	5.3	250	655	9.2	1.72
Iron α[d]	3.4×10^{4}	6.65	6.74	77.2	1183[e]	14.23	2.86	2.58	9.4	420	1074	7.07	3.63
Iron γ[d]	1.0×10^{-3}	−1.08	−0.76	48.0	1808[f]	5.78	3.56	2.57	9.4	420	1297	7.07	3.63

[a] Data from reference 1 unless otherwise noted.
[b] From Steigman, Shockley, and Nix, Phys. Rev. **56**, 13 (1939).
[c] From Raynor, Thomassen, and Rouse, Trans. Am. Soc. Met. **30**, 313 (1942).
[d] Data from reference 5.
[e] This is the transition temperature from α to γ as explained in text.
[f] This is the melting point of iron. A somewhat lower melting point for γ-iron proper would not influence the results appreciably.

[7] See, for example, W. A. Johnson, Trans. A.I.M.E. **143**, 107 (1941), also Table X of reference 2.

TABLE III. Data for intermetallic diffusion.*

System	Frequency factor, D_0, cm² sec.⁻¹	$\log\frac{D_0}{\nu d^2}$	$\log\frac{D_0}{\nu\lambda^2}$	Activation energy, E, kcal./mole	Melting point, T_m, °K	$\frac{E}{2.303RT_m}$	Lattice constant, d, cm $\times10^{-8}$	Interatomic distance, λ, cm $\times10^{-8}$	Debye frequency, ν, sec.⁻¹ $\times10^{13}$	Debye temp. θ	Av. T. of Exp. T_{Av} °K
Cd in Ag	4.9×10^{-5}	−2.19	−1.89	22.4	1234	3.98	4.08	2.88	4.5	215	1045
Cu in Ag	6.0×10^{-5}	−2.10	−1.80	24.8	1234	4.39	4.08	2.88	4.5	215	1045
In in Ag	7.3×10^{-5}	−2.01	−1.72	24.4	1234	4.33	4.08	2.88	4.5	215	1045
Sb in Ag	5.3×10^{-5}	−2.15	−1.85	21.7	1234	3.85	4.08	2.88	4.5	215	1045
Sn in Ag	7.8×10^{-5}	−1.98	−1.69	21.4	1234	3.80	4.08	2.88	4.5	215	1045
Cu in Al	8.4×10^{-3}	0.79	1.10	32.6	933	7.65	4.04	2.86	8.34	398	784
Cu in Au	1.1×10^{-3}	−0.77	−0.47	27.4	1336	4.48	4.07	2.87	3.8	170	1000
Al in Cu	1.8×10^{-2}	0.31	0.61	37.7	1357	6.07	3.61	2.55	6.6	315	1073
Be in Cu	2.3×10^{-4}	−1.57	−1.27	28.0	1357	4.52	3.61	2.55	6.6	315	1073
Cd in Cu	2.0×10^{-6}	−6.62	−6.33	8.0	1357	1.29	3.61	2.55	6.6	315	1073
Si in Cu	3.7×10^{-3}	0.63	0.94	40.0	1357	6.44	3.61	2.55	6.6	315	1073
Sn in Cu	1.1	2.12	2.42	45.0	1357	7.25	3.61	2.55	6.6	315	1073
Cu in Ni	1.0×10^{-3}	−1.0	−0.70	35.5	1728	4.50	3.51	2.48	7.85	375	1043
Cd in Pb	1.8×10^{-3}	−0.39	−0.08	15.4	600	5.62	4.94	3.48	1.8	88	483
Sn in Pb	4.0	2.96	3.27	26.2	600	9.53	4.94	3.48	1.8	88	538
Ni in Pb	6.6×10^{-1}	2.18	2.49	25.3	600	9.21	4.94	3.48	1.8	88	575
Tl in Pb	2.5×10^{-2}	0.76	1.07	19.4	600	7.06	4.94	3.48	1.8	88	515
Ag in 50:50 AgAu**	3.9×10^{-1}	1.75	2.05	44.7	1300	7.52	4.07	2.87	4.1	190	1111
Au in 50:50 AgAu**	1.2×10^{-1}	1.24	1.54	44.1	1300	7.41	4.07	2.87	4.1	190	1111

* Data from reference 1 unless otherwise stated. Only data obtained at very low concentrations and well below the solubility limits included. Lattice parameters and other physical constants refer to the solvent.
** Data from reference 9.

one, more subtle differences, specifically the relation of D_0 to lattice parameters, are extremely difficult to determine until a large amount of considerably more accurate experimental information is available.

The correlation is seen to be non-linear, except perhaps over the middle range, where a line of slope 1, indicated by the broken line, may be drawn. Over the linear range D_0 may be expressed as

$$D_0=\nu\lambda^2Ke^{E/RT_m} \qquad (2)$$

with $K=10^{-6}$. Since the linear form is valid only over a limited range, no physical significance can be assigned to the value of K.

It is to be noted that the anisotropy of Zn is properly taken care of in this correlation.* Furthermore, self-diffusion in α- and γ-iron also fits into this scheme provided the transition point between the two phases is accepted as the melting point of α-iron. This choice is not unreasonable in view of the fact that such a transition involves a great loosening of the lattice over the transition range of temperatures. Bismuth, which also shows anisotropic diffusion,[8] has been left out of Table II and Fig. 1. Available data on Bi are very approximate.[8] In the direction perpendicular to the C axis, a tremendously high D_0 was obtained ($\sim10^{47}$), which can be fitted in approximately with the present correlation. In the parallel direction a low D_0 is obtained ($\sim10^{-3}$) and the corresponding activation energy appears to be much too high. This point would fit very poorly on Fig. 1; but, because of the large scatter in the

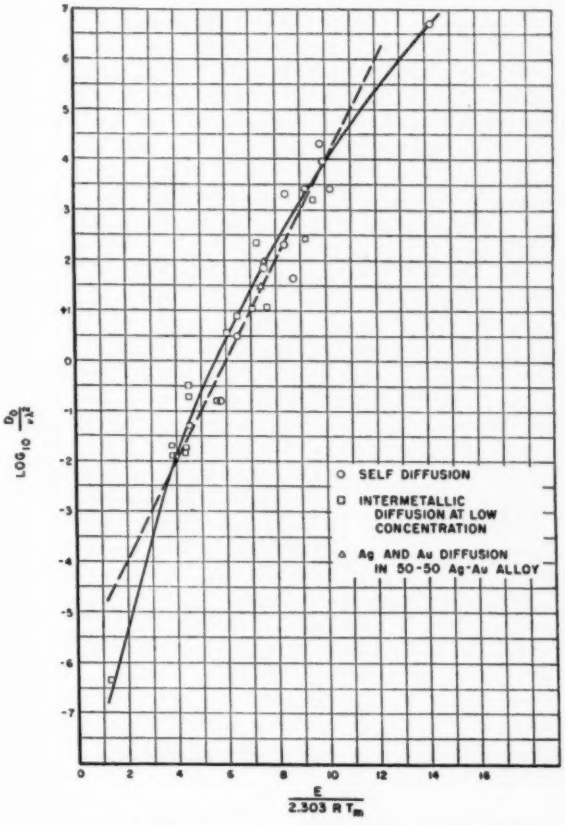


FIG. 1. Correlation of frequency factor with E/T_m.

* The Debye temperature given in Table II for Zn is an average value and, in the absence of better information, the same value was used for both directions. The characteristic temperature is unlikely to deviate by more than 30–40 percent from the average owing to anisotropy. This would have a negligible influence on the results since the frequency factors differ by more than a factor of 10.[3]

[8] W. Seith, Zeits. f. Elektrochemie **39**, 538 (1933). See also R. F. Mehl, Trans. A.I.M.E. **122**, 11 (1936).

data, it cannot be concluded that Bi is an exception to the correlation presented here. More experimental data on diffusion anisotropy as well as on the influence of phase changes (analogous to iron) would be very helpful.

It may be of interest also to note that Johnson's[9] values for the self-diffusion of silver and gold in a 50:50 AgAu alloy fall along the correlation curve.

III. THEORETICAL CONSIDERATIONS

It has been shown in the preceding sections that, empirically, the frequency factor, D_0, depends roughly exponentially on a definite fraction of the activation energy determined by the melting point. It is natural to look, therefore, for local melting, or disordering, as the elementary activation step. Such a proposal has been made by Mott[10] recently in connection with creep at low stresses. It appears that a similar picture may be applicable to diffusion. Let us assume that the elementary activation step is the disordering or melting of a small group of atoms by thermal fluctuations at a temperature T, which is below the melting point T_m, and let us assume that this small group of atoms may be treated thermodynamically. The free energy per atom necessary to produce this local melting, or disordering, may be written

$$\Delta F' = \Delta H' - T\Delta S' \tag{3}$$

with $\Delta F'$ equal to zero, at the melting point, T_m. Then

$$\Delta S' = \Delta H'/T_m. \tag{4}$$

The free energy of activation may be assumed to be proportional to $\Delta F'$ (the proportionality constant depending on the number of atoms involved),[10] i.e.,

$$F = \alpha\Delta F' = \alpha\Delta H' - T\alpha\Delta H'/T_m \tag{5}$$

or in the usual notation

$$F = E - TE/T_m, \tag{6}$$

where E/T_m represents the entropy of activation. Following reaction rate theory,[11] the expression for the diffusion constant D may be written

$$D = D_0\epsilon^{-E/RT} = \beta\epsilon^{E/RT_m}\epsilon^{-E/RT} \tag{7}$$

which is of the same form as the empirical equation valid over the linear range of Fig. 1. D_0 is identified with

$$\beta\epsilon^{E/RT_m},$$

where $\beta = K\nu\lambda^2$.

Thus, it is possible to lend some theoretical support to the empirical analysis presented in this paper.

It may be added that Eq. (7) has no physical meaning for $T > T_m$. At $T = T_m$

$$D = D_m = K\cdot\nu\lambda^2. \tag{8}$$

The analysis presented here states, therefore, that melting temperatures represent equivalent states for comparison of diffusion processes. Physically, it is far more reasonable to compare diffusion constants, and to relate them to lattice parameters, at the melting point rather than at $T = \infty$ as implied by the use of D_0. It would be desirable, of course to know accurately the lattice parameters at the melting point.

Another feature of the picture presented here may be pointed out qualitatively. Experimental results up to date indicate[2] that the activation energy for the diffusion of any metal A in dilute solution in another metal B is always less than the activation energy for self-diffusion in B. Johnson[9] has pointed out that the hole mechanism of diffusion is apparently unable to account for this trend and proposed that diffusion takes place via the motion of substituted atom-hole complexes. This suggestion has been criticized by Frenkel[12] as being improbable.

The picture of diffusion presented here can explain the above phenomenon qualitatively. The free energy of activation in Eq. (5) may be lowered either by decreasing α, i.e., the size of the region to be melted, or by decreasing the heat of fusion, $\Delta H'$. It seems unlikely that α would be changed appreciably. The heat of fusion, however, may be lowered considerably even in a dilute solution of A in B, since it is precisely around an atom of A that local melting must occur, and the concentration of A in the small group of B atoms may be quite high. Some change in the melting point is also to be expected modifying somewhat at the correlation of Fig. 1. The wide applicability of this correlation shows, however, that the important change must be in the activation energy. No quantitative results can be given at this time.

With the importance of melting characteristics, as related to diffusion, empirically established, it seems likely that considerable further theoretical progress may be made by a careful investigation of melting and local disorder phenomena. Modern theories of melting and fluctuation phenomena should form a convenient starting point.

Note added in proof: Since this paper was written, the author's attention has been called to a recent paper by C. F. Birchenall and R. F. Mehl, Trans. A.I.M.E., 188, 144 (1950), in which new data are given for self-diffusion in α- and γ-iron which supersede entirely their previous results (Reference 5). These new data fit into the correlation scheme presented here and do not change the conclusions given in the text. As a matter of fact, since the activation energies for the two phases are now practically identical, the explanation in terms of melting points may acquire added significance.

[9] W. A. Johnson, Trans. A.I.M.E. **147**, 331 (1942).
[10] N. F. Mott, Proc. Phys. Soc. London **60**, 391 (1948).
[11] S. Glasstone, K. J. Laidler, and H. Eyring, *The Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, 1941), pp. 519, 538.
[12] J. Frenkel, *Kinetic Theory of Liquids* (Oxford University Press, Oxford, 1946), pp. 27–36.

Letters to the Editor

Thorium Sulfide as a Thermionic Emitter

T. E. Hanley
Naval Research Laboratory, Washington, D. C.
June 28, 1950

IN a recently declassified AEC report[1] it is stated that sulfides of thorium, particularly ThS, have high resistance to thermal shock and are "more resistant toward metals than the oxides." In addition, ThS has an electrical conductivity of $10^{-1}\Omega$ cm and "is as non-volatile as the most refractory oxides." In view of these characteristics it was of interest to investigate the possibility of using ThS as a thermionic emitter.

A sample of this material was obtained through the AEC. A coating bath was set up in order to obtain uniform coatings. The bath was prepared by ball-milling (Pyrex balls in Pyrex container) 5 g of ThS with a mixture of 50 ml of methanol and 50 ml of acetone. This combination when used as a coating bath with a nickel anode would put a 0.001-in. coating on a 0.0005-in. wire in 10 sec. at 50 v and an initial current density of 40 ma/cm^2.

The temperature scale for ThS was obtained by heating a ThS-coated 0.002-in. molybdenum cylinder (o.d. = 0.125 in., L = 3.5 in.) in vacuum. The brightness temperature was measured on the outer coated surface, while the true temperature was obtained by sighting the optical pyrometer on a 0.30-in. hole in the cylinder. The measurements give a spectral emissivity of 0.4 for ThS.

The thermionic emission data were taken using coated 0.005-in. tungsten wires in a guard ring diode structure. The Richardson constants of the material are approximately Φ=3.4 ev and A = 100 amp./cm^2/deg.^{2}. In the vicinity of 1500°C Br the emission from ThS is only one-fourth as great as that from thoria (ThO_2).

Some measurements were also made to determine the rate of evaporation of this material by measuring the loss of weight from a uniformly coated tungsten strip which was held at a known temperature for a given length of time in vacuum. These measurements indicate a vapor pressure of 10^{-3} mm Hg at 1870°C Br whereas the previously mentioned report[1] gives the vapor pressure as less than 10^{-3} mm Hg at 2200°C. Perhaps comparison of these data is not in order, since the conditions of the measurements for the values given in the AEC report are not stated.

[1] Brewer, Bromley, Giles, and Lofgren, "The preparation and properties of refractory sulfides," University of California (November, 1945).

On the Fredrickson-Eyring Theory of the Mechanical Behavior of Metals*

J. Fleeman† and G. J. Dienes
Atomic Energy Research Department, North American Aviation, Inc., Downey, California
August 28, 1950

IT is generally felt that a unified theory of the mechanical behavior of metals must be based upon a dynamical theory since it is well known that the results of a stress-strain experiment depend markedly upon the time scale of the experiment. The first serious attempt to formulate a dynamical theory of the behavior of metals is due to Fredrickson and Eyring.[1] We shall refer to this paper as FE. It is the object of this letter to discuss some of the implications of their theory.

FE have used the method of the statistical rate theory of processes to formulate a theory of the mechanical behavior of metals. The behavior of the fundamental unit of dashpot and spring is determined by the differential equation

$$d\epsilon/dt=(1/g_i)(d\sigma/dt)+k_i\sinh\alpha_i\sigma, \tag{1}$$

where σ and ϵ are the stress and strain, respectively, g_i is the spring constant of the ith element and, k_i and α_i are the viscosity parameters determined by way of a quantity β_i. The first term on the right-hand side of the equation describes the elastic properties of metals according to Hooke's law; the last term describes the flow properties according to the Eyring theory of viscosity. As a model used to interpret a stress-strain curve, FE use the combination of springs g_1, g_2, and g_3 and dashpot elements β_1 and β_2 in the form of Maxwell elements coupled in parallel. The spring g_1 and dashpot β_1 combine to form unit one, unit two consists of the spring g_2 and dashpot β_2, unit three is the spring element g_3 alone. FE have applied the results of their analysis to determine the equation for a loading curve obtained with a constant rate of straining.

To test the generality of this dynamical theory we have used Eq. (1) to calculate an unloading curve and (what turns out to be more interesting) a reloading curve calculated from the assumption that the material is reloaded immediately after unloading. The problem of strain relaxation, or recovery, at zero load is discussed separately under several simplifying assumptions.

With the test material loaded to a stress $\sigma_0^{(1)}$ at a strain $\epsilon_0^{(1)}$ under the conditions of a constant strain rate $\dot{\epsilon}$ the unloading curve can be evaluated by solving Eq. (1) subject to the initial condition $\sigma^{(1)}=\sigma_0^{(1)}$ at $t=0$. We get

$$\sigma^{(1)}=\sigma_0^{(1)}-\frac{1}{\alpha}\ln\left\{\frac{1+\left[1+\frac{\exp(\alpha\sigma_0^{(1)})}{\beta}\right]\tanh\left[\frac{\alpha g}{2}(\epsilon_0^{(1)}-\epsilon)\right]}{1-\tanh(\alpha g/2)(\epsilon_0^{(1)}-\epsilon)}\right\}. \tag{2}$$

The unloading curve for each unit of the Maxwell representation is calculated from Eq. (2) to a value of zero stress for the entire model. It turns out that after unloading the first unit is in a state of compression while the second and third units are in extension.

Similarly, the reloading curve can be determined from Eq. (1) using the initial condition that at $t=0$, $\sigma=\sigma_0^{(2)}$, where $\sigma_0^{(2)}$ can be positive or negative depending upon whether it refers to unit one or two. We get

$$\sigma^{(2)}=\sigma_0^{(2)}+\frac{1}{\alpha}\ln\left\{\frac{1+\left[1+\frac{\exp(-\alpha\sigma_0^{(2)})}{\beta}\right]\tanh\left[\frac{\alpha g}{2}(\epsilon-\epsilon_0^{(2)})\right]}{1-\left[1-\frac{\exp(\alpha\sigma_0^{(2)})}{\beta}\right]\tanh\left[\frac{\alpha g}{2}(\epsilon-\epsilon_0^{(2)})\right]}\right\}. \tag{3}$$

We have calculated the unloading and reloading stress-strain curves for SAE 1020 steel according to Eqs. (2) and (3), respectively,

1. at $T=-70$°C and a strain rate of 5×10^{-5} sec.$^{-1}$,
2. at $T=-70$°C and a strain rate of 500×10^{-5} sec.$^{-1}$,
3. at $T=100$°C and a strain rate of 5×10^{-5} sec.$^{-1}$,

where we have assumed the material has been loaded to a value of true strain equal to 0.5 and then unloaded.

The results of these calculations are shown in Fig. 1. It will be noted that in all cases the reloading curve rises appreciably above the initial loading curve for equivalent values of strain. It appears that the model predicts too much strain hardening.[2]

It is important to determine the magnitude and rate of strain recovery to be expected from the FE theory. Such a calculation is to be carried out via Voigt-type of mechanical model representation. However, as a result of the non-linear character of Eq. (1), the equations which arise in the transformation of a Maxwell representation to a Voigt representation are too complex to interpret in a simple manner. In order to make an estimate of the relaxation time we have replaced the non-linear viscous element in the Maxwell representation by an equivalent Newtonian viscous element η. If we choose the viscosity η to be given by the equation

$$\eta=1/\alpha k, \tag{4}$$

then the relaxation time which is calculated will always be a time that is greater than any relaxation time obtained from the FE theory. Furthermore, it can be shown that another simplification which may be made, without appreciably influencing the results is the following: letting $g_1=0$ in the Maxwell representation renders the equivalent Voigt representation, consisting of springs

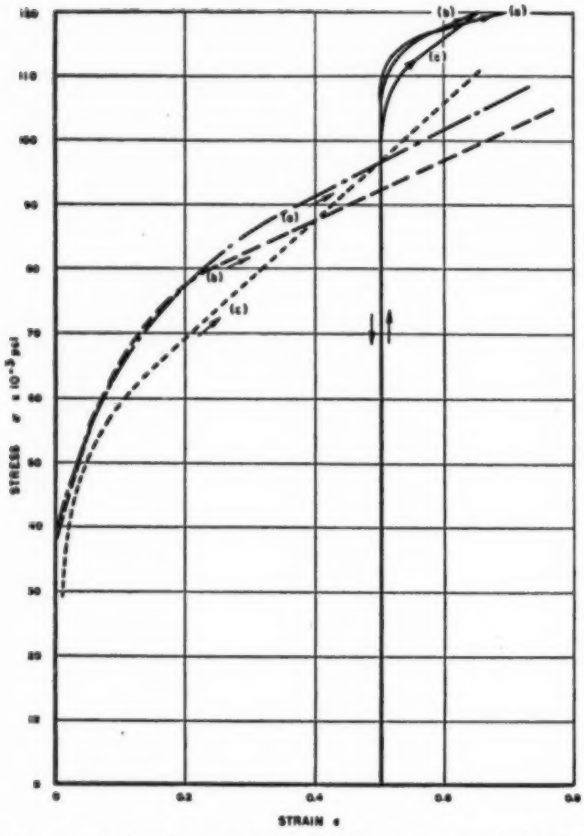


FIG. 1. The dotted curves are the loading curves as calculated by Fredrickson and Eyring. Curve (a) computed for $T=-70°C$ and $\dot{\epsilon}=500\times10^{-6}$ sec.$^{-1}$, curve (b) computed for $T=-70°C$ and $\dot{\epsilon}=5\times10^{-6}$ sec.$^{-1}$, curve (c) computed for $T=100°C$ and $\dot{\epsilon}=5\times10^{-6}$ sec.$^{-1}$. The solid curves are the unloading and reloading curves computed from a strain $\epsilon=0.5$ as discussed in the text.

g_1' and g_2' and Newtonian viscous elements η_1' and η_2' easily manageable.

Using the method outlined, we have calculated the maximum relaxation time for SAE 1020 steel at $T=-70°C$ and $\dot{\epsilon}=5\times10^{-6}$ sec.$^{-1}$ to be 2.6 days and 2.2 hours for the two units which appear in the Voigt representation. It is evident that strain relaxation is going to take place at a rate far greater than experimentally expected for steel in a temperature range where plastic deformation should be essentially permanent.[3]

Thus, upon the basis of calculations presented here, we conclude that the FE dynamical theory of the mechanical behavior of metals has inherent within it the following characteristics:

1. It predicts too much strain hardening.
2. It predicts a relaxation time for deformation recovery which is much too small and consequently it fails to account for the permanent deformation in steel in the temperature range $-70°C$ to $100°C$.

* This letter is based on studies conducted for the AEC under Contract AT-11-1-GEN-8.
† Now at the National Bureau of Standards, Washington, D. C.
[1] J. W. Fredrickson and H. Eyring, *Statistical Rate Theory of Metals*, Metals Tech. TP 2423, August, 1948.
[2] See for example: F. Seitz, *The Physics of Metals* (McGraw-Hill Book Company, Inc. New York, 1943), Fig. 46, p. 72; Fig. 51, p. 83.
[3] A. M. Freudenthal and M. Reiner, J. App. Mech. 15, 265 (1948).

Base-Metal Effects in Thoria-Coated Filaments

HERBERT NELSON
Tube Department, Radio Corporation of America, Harrison, New Jersey
August 11, 1950

THE thoria-coated filament has shown promise of meeting the demands for a thermionic source supplying long-life emission at high current densities and voltages better than does either the barium-strontium oxide-coated cathode or the thoriated tungsten filament. Investigation of the properties of this new type filament has therefore appeared fruitful and has been carried out in various laboratories of the RCA organization. The purpose of the present communication is to summarize briefly some of the results obtained in the part of this investigation which has had to do with base-metal effects. Adherence of the thoria to the metal base, ease of thermionic activation, and level of thermionic emission reached have been studied as a function of base material.

With regard to adherence, a problem is created as a result of the rather large difference in thermal expansion between thoria and the refractory metals suitable for base material. The coefficient of thermal expansion for thoria[1] is 11.3×10^{-6}, while for W, Mo, Ta, and Pt, the values are 4.2, 4.9, 7.0, and 9.0×10^{-6}, respectively. On the basis of these figures it is to be expected that adherence is best with Pt and poorest with W. Experimental work has shown the foregoing to be true. Adherence was found poor with W, fair with Mo, good with Ta, and excellent with Pt. With Ta, layers of thoria as thick as 3 mils (45 mg/cm^2) were found to adhere to the end of 3000-to-4000-hour life tests conducted in diodes. With W, on the other hand, even very thin layers (8 mg/cm^2) did tend to chip off at the very beginning of life tests. With Mo, chipping difficulties were also encountered though filaments with as much as 20 mg/cm^2 of thoria were operated successfully on life for as long as 2000 to 3000 hours.

Though, as suggested above, the better match in thermal expansion may account for the superior adherence to Pt and Ta, it is also possible that the relatively high ductility of these metals at operating temperatures may be a contributing factor.

Since it seems reasonable to expect that for low work function, thoria, like BaSrO, depends upon a stoichiometric excess of metal in the crystal lattice, tests were made to determine whether the base metal, as a reducing agent, has an effect upon the ease of activation of thoria-coated filaments. For this purpose four tubes were constructed identical in all respects except that two were made with a base material of Ta while the other two were made with Pt. The tubes were exhausted and processed in identical fashion, the electron emission being determined as a function of the filament temperature during the initial glowing. The results obtained are shown in Fig. 1, curves *A* and *B* expressing the temperature-electron emission relation for thoria on Ta and Pt, respectively. The fact that high thermionic activity is reached at

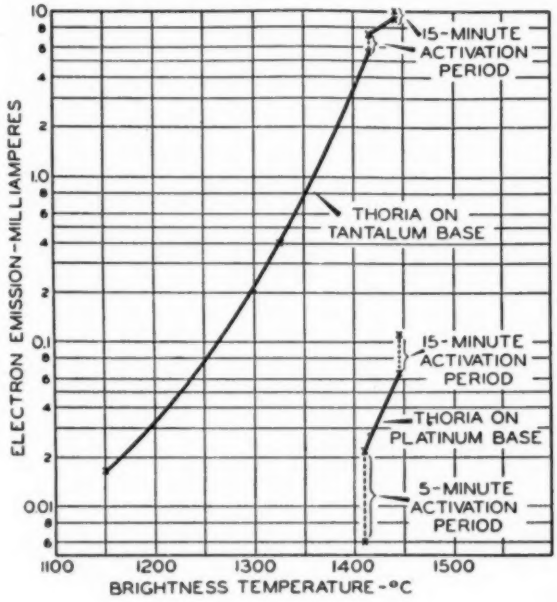


FIG. 1.

a greatly lower temperature with Ta, a good reducing agent, than with Pt, a poor reducing agent, is consistent with the view that a low work function of thoria is associated with a stoichiometric excess of metal in the crystal lattice.

When Mo, W, or Ta is used as base material, roughly the same maximum space-charge-limited emission, about 5 amp./cm², is obtained at 1900°K B.T. At lower temperatures, however, higher values are invariably obtained with Mo than with Ta. At 1800°K B.T., for instance, the average emission with Mo was found to be equal to 3.1 amp./cm², while the average with Ta was found to be 2.0 amp./cm². No explanation has as yet been found for this superiority of Mo at low temperatures.

Acknowledgment is made to the following members of the RCA organization who have made fundamental contributions to the development of the thoria-coated filament: Dr. Lloyd Garner, Dr. Philip Smith, and Mr. M. N. Fredenburgh.

[1] A. Eisenstein, J. App. Phys. **17**, 443 (1946).

Erratum: The Mechanical Properties of Glass

[J. App. Phys. **13**, 623–634 (1942)]

F. W. PRESTON
The Preston Laboratories, Butler, Pennsylvania

ON p. 631 there is an error in the caption for Fig. 5. The limiting velocity of the propagation of cracks should be 5080 ft./sec., rather than 3 miles/sec. as stated.

A similar error occurs on p. 624, in which the velocity of sound in glass should be approximately 1 mile/sec., instead of 3 miles/sec.

Scheme for Analyzing Capillary Measurements on Non-Newtonian Liquids

ORRIN H. CLARK AND M. L. DEUTSCH
Socony-Vacuum Laboratories, Research and Development Department, Paulsboro, New Jersey

August 11, 1950

IT has kindly been pointed out by Mr. A. B. Bestul that the formulas for analyzing non-Newtonian capillary flow given in our recent paper, "Scheme for analyzing capillary measurements on non-Newtonian liquids,"[1] were developed by B. Rabinowitsch, "Über die Viskosität und Elastizität von Solen,"[2] and have since been referred to in various other journals. The absence of these useful equations from such a book as Green's *Industrial Rheology and Rheological Structures* is surprising and led to the regrettable submission by us of their derivation as original material.

[1] Orrin H. Clark and M. L. Deutsch, J. App. Phys. **21**, 713 (1950).
[2] B. Rabinowitsch, Zeits. f. physik. Chemie **A145**, 1–26 (1929).

Original Authorship of Formulas Given in Clark and Deutsch's Paper

M. REINER
Institute of Technology, Haifa, Israel
AND
G. W. SCOTT BLAIR
National Institute for Research in Dairying, University of Reading, England

September 14, 1950

THE fact that the paper by Clark and Deutsch in the July issue of your Journal (p. 713) was passed by your referee as describing "a powerful new scheme" shows that rheological literature is not as widely known as it should be. May we, therefore, as authors of textbooks on rheology in which the relevant papers are quoted (Scott Blair, 1938; Reiner, 1949) be permitted to point out that this method is due to Weissenberg and was first applied by B. Rabinowitsch (1929) for the case of the capillary viscometer, followed by Mooney (1931), who adapted it to the rotating cylinder instrument? Hersey (1932) named it the "differentiation method" because it applies differentiation to the observed consistency curve in order to find the rheological equation of the material, in contradistinction to the "integration method," where a rheological equation is assumed and the consistency curve derived by integration.

We do not doubt that Messrs. Clark and Deutsch derived their Eq. (12) independently, but it will be seen that it is identical with Eq. (9) of Hersey's paper.

Application of Liouville's Approximation to the Blind Navigation Problem

J. J. GILVARRY AND S. H. BROWNE*
The Rand Corporation, Inc., Santa Monica, California

September 8, 1950

IN a recent paper[1] in this journal, it has been shown that the problem of determining a vehicle's position vector $\mathbf{r}$ from the reading $\mathbf{b}(t)$ of an accelerometer as a function of time t is equivalent to solving the differential equation

$$(d^2\mathbf{r}/dt^2)+\omega_0^2(R_0/R)^3(R_0\mathbf{k}+\mathbf{r})=\mathbf{b}(t), \tag{1}$$

where R (with initial value R_0) is the vehicle's radial distance from the earth's center, $\mathbf{k}$ is a radial unit vector at the origin, and $\omega_0=(g_0/R_0)^{\frac{1}{2}}$ if g_0 is the gravitational acceleration at the initial point of motion. A case not treated in that discussion presents itself when the vehicle's altitude $h=R-R_0$ relative to the earth is available as a function $h(t)$ of time from continuous measurements made by an altimeter (of radar or atmospheric-pressure type). The quantity $h(t)$ is then an independently measured integral of the motion (thus the problem of determining position coordinates is actually overdetermined). In this case, Eq. (1) becomes

$$(d^2\mathbf{r}/dt^2)+\omega_0^2\Gamma(t)\mathbf{r}=\mathbf{b}(t)-g_0\Gamma(t)\mathbf{k}, \tag{2}$$

where

$$\Gamma(t)=[1+h(t)/R_0]^{-2}\sim 1-3h(t)/R_0, \tag{3}$$

and the approximation indicated in (3) is valid for $h\ll R_0$. Equation (2) is linear in the coordinates and, from the standpoint of computer design, presents definite advantages over the nonlinear Eq. (1).

Any numerical method of solving a linear differential equation of second order will serve as a computational method to determine $\mathbf{r}(t)$ from a given $\mathbf{b}(t)$ and $h(t)$, or, when suitably mechanized, will serve as the basis of a computer design. For numerical computation, however, the majority of such methods are not well adapted to equations of the type (2), where one term (the $\mathbf{g}$-term) is a slowly varying function, but another term (the $\mathbf{b}$-term) may show rapid variations in time. When applicable, though, Liouville's approximation[2] is well suited to problems of this kind, and this letter indicates briefly its application to (2). The approximation in question is known as the Wentzel-Kramers-Brillouin-Jeffries approximation in its application as an asymptotic approximation in quantum mechanics; more generally, it appears in connection with transformations and asymptotic developments of Sturm-Liouville systems.[3] It is ascribed to Liouville by Schelkunoff.[4]

The quantity Γ (which is always positive for $|h|<R_0$) is approximately unity except for the perturbing parameter $h(t)$. Under the transformation

$$t'=\int_0^t \Gamma^{1/2}(\sigma)d\sigma, \quad \mathbf{r}'=\mathbf{r}\Gamma^{1/4}(t), \tag{4}$$

Eq. (2) yields the approximation

$$(d^2\mathbf{r}'/dt'^2)+\omega_0^2\mathbf{r}'=\Gamma^{-3/4}(t)\mathbf{b}(t)-g_0\Gamma^{1/4}(t)\mathbf{k}, \tag{5}$$

provided the first and second derivatives of $h(t)$ are sufficiently small so that

$$\Gamma^{-3/4}(d^2\Gamma^{-1/4}/dt^2)\ll\omega_0^2, \tag{6}$$

and hence the term $\Gamma^{-3/4}(d^2\Gamma^{-1/4}/dt^2)\mathbf{r}'$ can be neglected in (5). When the right-hand side of Eq. (5) is considered a function $\mathbf{b}'$

of t', the equation is linear with constant coefficients in the reduced variables t', r'. The explicit solution for r then follows directly as

$$r=(v_0/\omega_0\Gamma^{1/4})\sin\left[\omega_0\int_0^t\Gamma^{1/2}(\sigma)d\sigma\right]+\frac{1}{\omega_0\Gamma^{1/4}}$$
$$\times\int_0^t\left[\frac{b(\tau)}{\Gamma^{1/4}(\tau)}-g_0\Gamma^{3/4}(\tau)k\right]\sin\left[\omega_0\left(\int_0^t\Gamma^{1/2}(\sigma)d\sigma-\tau\right)\right]d\tau, \quad (7)$$

under the initial conditions $r_0=0$ and $(dr/dt)_0=v_0$. In actual application, Γ in (7) can usually be replaced by the approximation indicated in (3).

Equation (7) yields a convenient solution for numerical computation, since the solution is reduced to quadratures which can be carried out by the trapezoid or Simpson's rule (the sine function under the integral sign must be expanded by the addition theorem). To apply the method over an extended interval (over which the effect of the neglected term in Eq. (5) may be significant), one can use (7) as the step solution in the method of "linear continuation" described in the previous paper.[1] Alternatively, Schelkunoff's "wave perturbation" method[4,5] (with a slight modification for the presence of a forcing function) may be used; this method is based on Liouville's approximation, but defines the final solution as the limit of an iterative process over the over-all interval.

* Now at North American Aviation, Inc., Los Angeles, California.
[1] Gilvarry, Browne, and Williams, J. App. Phys. 21, 753 (1950).
[3] J. Liouville, J. de Math. 1, 253 (1836); 2, 16 (1837); 2, 418 (1837).
[4] R. Courant and D. Hilbert, *Methoden der mathematischen Physik* (Verlag. Julius Springer, Berlin, 1931), Vol. I, pp. 250, 290.
[4] S. A. Schelkunoff, Q. App. Math. 3, 348 (1945).
[5] M. C. Gray and S. A. Schelkunoff, Bell Syst. Tech. J. 27, 350 (1948).

Natural Crystal Greenockite (Cadmium Sulfide) as an X-Ray Detector

E. L. Criscuolo and D. T. O'Connor
U. S. Naval Ordnance Laboratory, Silver Spring, Maryland
September 8, 1950

RECENTLY,[1] artificially produced cadmium sulfide crystals have been described as sensitive x-ray detectors. Previous work described in the literature[2] refers to artificial cadmium sulfide crystals.

Natural greenockite crystal fragments of the order of 1 mm³ have been observed to pass direct currents up to 1×10^{-6} amp. in a 100-kv x-ray beam of an intensity of 4 roentgens per minute. This rough measurement was made using a micro-ammeter and a collecting battery voltage of 22 volts. Four samples gave similar results; they were light-sensitive and non-luminescent.

The mineral specimen was obtained from the Smithsonian Institute through the courtesy of Dr. W. F. Foshag, Chief of the Division of Mineralogy and Petrology, and Dr. R. Kellogg, Director.

[1] Rudolf Frerichs, J. App. Phys. 21, 312–317 (1950).
[2] Robert Hofstadter, Nucleonics (April, 1949); (May, 1949).

Measurements on the Diffusion of Interstitial Atoms in B.C.C. Lattices*

C. A. Wert
Institute for the Study of Metals, University of Chicago, Chicago, Illinois
August 14, 1950

THE alloys of interstitial dissolved atoms in the b.c.c. lattices offer the possibility that one can make extremely accurate measurements of the rate of diffusion of the solute atoms. This is so because the high temperature bulk diffusion measured by standard metallurgical methods can be extended to lower temperatures by measurements made on an atomistic scale by use of relaxation measurements. The diffusion of C in α-Fe offers an example of this technique. Stanley[1] measured the rate of diffusion in the range 500°C to 700°C. The author[2] measured diffusion in the range −35°C to +200°C. Together the data extend the knowledge of the diffusion coefficient over some 14 cycles of 10; this range is somewhat higher than has been done for any other system.

In an effort to extend this type of measurement to other systems, the author has made measurements of diffusion of C in Ta and Cb and of N in α-Fe. These measurements were made by two manifestations of the anelastic relaxation effects; the internal friction and the elastic aftereffect. The methods used are those described earlier.[2] The data are shown in Fig. 1. Here the diffusion

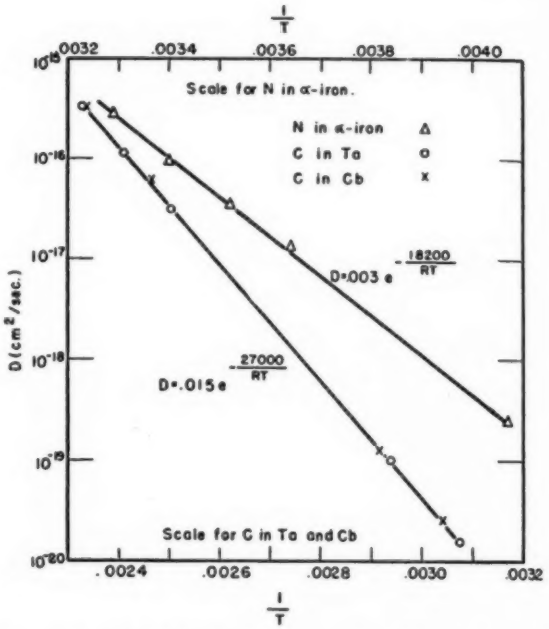


FIG. 1. Diffusion coefficient for C in Ta and Cb and N in α-Fe *vs.* $1/T$.

of N in α-iron and of C in Ta and Cb are seen to satisfy the well-known relation $D=D_0\exp(-H/RT)$. These data, of course, do not have the range quoted above for C in α-Fe; further measurements at much higher temperatures are needed for really accurate determination of D_0 and H.

An attempt to examine the various factors which make up D_0 has been made by Wert and Zener.[3] They expressed D_0 as

$$D_0=n\alpha a^2\nu e^{\Delta S/R}, \quad (1)$$

where n, α, and a^2 are geometrical constants for a given lattice, ν is the frequency of oscillation of a solute atom in an interstitial site, and ΔS an "entropy of activation." They found that if all of the activation energy for diffusion went into straining the lattice, then

$$\Delta S\approx -H(d\ln\mu/dT),$$

where μ is the shear modulus of the lattice. For simplicity one can rewrite this expression by defining $\beta=d(\mu/\mu_0)/d(T/T_m)$ obtaining

$$\Delta S\approx\beta H/T_m. \quad (2)$$

In this expression μ_0 is the modulus at $T=0$ and T_m is the melting temperature. β may be obtained for a large number of materials from some work of Köster.[4] For Ta and α-Fe the values of β are 0.40 and 0.43, respectively.

It follows from Eq. (1) that once D_0 is known from experiment, ΔS may be calculated. It is also true that if all the energy of activation goes into lattice strain energy, then ΔS is given by Eq. (2). A comparison of $\Delta S/R$ calculated by the two methods is given in Fig. 2. Here are shown data for C in α-Fe and Ta, N in

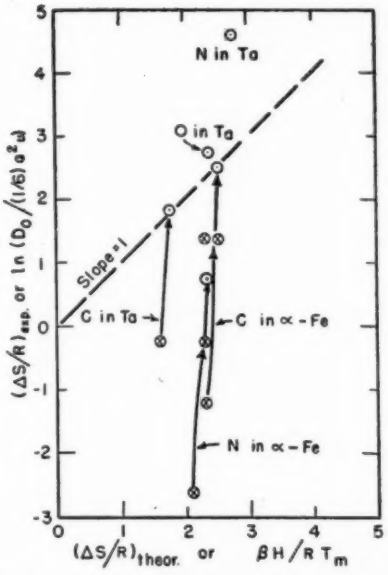


FIG. 2. $(\Delta S/R)$ as computed from Eq. (1) using diffusion experimental data *vs.* $\Delta S/R$ as computed from Eq. (2). For each material successive points are taken from the literature and the present letter as follows.

Oxygen in Ta	T. Ke, Phys. Rev. **74**, 9 (1948).
Nitrogen in Ta	T. Ke, Phys. Rev. **74**, 914 (1948).
Carbon in Ta	T. Ke, Phys. Rev. **74**, 9 (1948); C. A. Wert (present letter).
Carbon in α-Fe	J. Snoek, Physica **8**, 711 (1941); J. K. Stanley, Trans. A.I.M.E. **185**, 752 (1949); C. Wert and C. Zener, Phys. Rev. **76**, 1169 (1949); C. Wert, Phys. Rev. (to be published)
N in α-Fe	J. Snoek, Physica **8**, 711 (1941); C. Wert and C. Zener, Phys. Rev. **76**, 1169 (1949); C. Wert (present letter).

α-Fe and Ta, and O in Ta. The open circles represent the best data available at present; the other circles represent earlier data as given in the caption to Fig. 2. If $\Delta S/R$ computed by Eqs. (1) and (2) were identical, then all the points should lie on the dashed line. The trend of the measurements toward better agreement between $\Delta S/R$, as computed by the two methods, is quite apparent. This trend toward better agreement presumably results from the better values of D_0 obtainable by extension of the measurements over a wider range of temperature.

* This research was supported in part by the ONR (Contract No. N-6ori-20-IV, NR 019 302).

[1] J. K. Stanley, Trans. A.I.M.E. **185**, 752 (1949).
[2] C. Wert, Phys. Rev. (to be published).
[3] C. Wert and C. Zener, Phys. Rev. **76**, 1169 (1949).
[4] W. Köster, Zeits. f. Metallkunde **39**, 1 (1948).

A Remark about Standardization and "A Dimensional Analysis of Metal Cutting" *

ANDRÉ MARTINOT-LAGARDE
Associate Professor of Mechanics of Fluids, Lille University, France
September 8, 1950

ANALOGOUS non-dimensional variables appear in both metal cutting and mechanics of fluids. Therefore, it seems convenient to choose exactly the same non-dimensional variables. For instance in mechanics of fluids $c\rho Vl/K$ (c specific heat, ρ mass density, V velocity, l length, K thermal conductivity) is called the "Péclet number." This note is to propose writing the same number for metal cutting, instead of its inverse, used by Drucker and Ekstein, and using the same designation also. The velocity indeed is the variable which varies the most frequently, and it is convenient that every non-dimensional variable be proportional to such a quantity.

* D. C. Drucker and H. Ekstein, J. App. Phys. **21**, 104–107 (1950).

Tearline Patterns in Ferrochromium

C. A. ZAPFFE AND F. K. LANDGRAF
Baltimore, Maryland
August 14, 1950

IN the course of studying fractographic patterns,* special formations have frequently been observed within the fracture traverse which are believed to relate to conditions of original crystal growth. Some suggestively similar patterns, however, are discussed from another viewpoint in the recent paper by Kies, Sullivan, and Irwin in this journal.[1] Their contribution is an important advance in the interpretation of fracture patterns, and specifically those previously classified as Type II expressing transient conformations of the fracturing stress.[2] It also focuses attention upon a principal current problem in the interpretation of certain fracture patterns—the distinguishing between Types I and II. Many fractographs, without possible question, reveal structural features of the solid; but others clearly refer only to stress. Because the correct identification often has important implications for theories of the solid state, several complex examples are submitted here, taken from a study of chromium-rich iron.

FIG. 1. Fractograph of ferrochromium containing 70 percent Cr and 0.05 percent C, showing a transition from transgranular to intergranular fracture traverse.

In Fig. 1, transgranular and intergranular fractures lie within the same general plane of focus. To the left of the diagonal boundary, fracture has followed a cleavage plane of the icositetrahedron {112}, displaying a relatively flat traverse with traces of intersecting planes which can be identified as also belonging to the icositetrahedral form. The traverse then shifts from a planar to a banal path along grain or lineage boundaries.

In Fig. 2 some of the fanlike forms visible on the banal traverse of the previous figure are shown in more detail. The condition here is particularly interesting because of an irregular boundary roughly bisecting the field. Patterns somewhat similar to these, as has been shown by Kies, Sullivan, and Irwin, can be obtained

FIG. 2. Intergranular surfacial detail in fractured 70 percent ferrochromium containing 0.17 percent C.

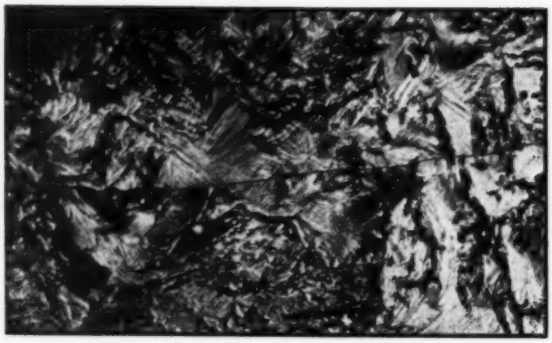


FIG. 3. Fractograph of ferrochromium containing 0.75 percent nitrogen, showing a boundary [lightly retouched for reproduction] which sharply separates the upper and lower halves of the field.

by special stress conformations during fracture of amorphous or rheologically isotropic bodies. However, solid metals show pronounced directionalism, both from crystal perfection and from crystal imperfection; and according to the theory advanced earlier in this study,[3] the present authors prefer the interpretation that these formations relate to original crystal growth in which clusters of molecules—"micellès"—have assumed a common orientation to provide the single grain, but have carried into the solid state certain aspects of their original micellar form and a general imperfection pattern of their cumulative mismatching. It is suggested that the grain under observation in Fig. 2 has grown from the lower right as the result of a commonly oriented accumulation of two major lineages and several minor ones. This growth then possibly halted along the observed boundary during solidification, beginning again with renewed nucleation, as indicated in the pattern. On the other hand, descriptions given by Kies, Sullivan, and Irwin are impressive in suggesting that at least a portion of this pattern might be Type II, the indicated nucleations being nucleations of fracture with or without further impression of imperfection structure.

In Fig. 3 a particularly puzzling pattern appears in which a boundary of the type in the previous figure delineates with sharp contrasts the upper and lower portions of the field. It is possible that elaborately complex stress formations during fracture approached simultaneously from above and below, meeting at the indicated boundary to form the observed pattern.

On the other hand, it is submitted as the preferred interpretation here that an unusual growth condition is being observed in which dendritic formations, separately nucleated in the liquid phase and growing with lineage characteristics, have approached from solidification directions above and below in the fractograph, the two masses mutually exhausting the remanent liquid phase as solidification progressed, producing thereby both the boundary and the branching fractographic forms.

* Investigation conducted in the senior author's Laboratory under contract with the ONR.

[1] Kies, Sullivan, and Irwin, J. App. Phys. **21**, 716–720 (1950).
[2] C. A. Zapffe and C. O. Worden, Acta Cryst. **2**, Part 6, 377–82 (1949).
[3] C. A. Zapffe, Am. Soc. Metals Trans. **42**, 387–98 (1950).

Note on "Interference of Growing Spherical Precipitate Particles"

HERSHEL MARKOVITZ
Mellon Institute of Industrial Research, University of Pittsburgh, Pittsburgh, Pennsylvania

July 28, 1950

IN a recent article Wert and Zener[1] derived the differential equation

$$\tau dW/dt=(3/2)(1-W)W^{1/3}$$

to which a solution was desired with the condition that at $t=0$, $W=0$. An analytical integration of this equation is

$$\frac{t}{\tau}=\frac{2}{3}\left[\frac{1}{2}\ln\left(\frac{1+x+x^2}{(1-x)^2}\right)+\sqrt{3}\tan^{-1}\left(\frac{-x\sqrt{3}}{2+x}\right)\right],$$

where $W=x^3$. If various values of W are substituted, corresponding values of t can be obtained. As far as one can tell from their Fig. 3, this solution seems to lead to somewhat better agreement with the experimental results than the numerical integration performed by the authors.

[1] C. Wert and C. Zener, J. App. Phys. **21**, 5 (1950).